ANNÉE 2018

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Bretagne Loire

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Informatique

Ecole doctorale MATHSTIC

présentée par

Jiating LUO

Préparée à l'unité de recherche UMR6074 IRISA
Institut de recherche en informatique et systèmes aléatoires - CAIRN
Université de Rennes 1

Architectural and Protocol Exploration for 3D Optical Network-on-Chip

Thèse soutenue à Lannion
le 11 juillet 2018

devant le jury composé de :

Olivier ROMAIN
Professeur à l'Université de Cergy-Pontoise/ *rapporteur*

Cécile BELLEUDY
Maître de Conférences - HDR à l'Université Nice Sophia Antipolis/ *rapportrice*

Jean-Philippe DIGUET
Directeur de recherche CNRS à l'Université Bretagne-Sud / *examinateur*

Laurent VIVIEN
Directeur de recherche CNRS à l'Université Paris-Sud/ *examinateur*

Camel TANOUGAST
Professeur à l'Université de Lorraine/*examinateur*

Daniel CHILLET
Professeur à l'Université de Rennes 1/ *directeur de thèse*

Cédric KILLIAN
Maître de Conférences à l'Université de Rennes 1/ *co-encadrant de thèse*

Sébastien LE BEUX
Maître de Conférences - HDR à l'École Centrale de Lyon / *co-encadrant de thèse*

Acknowledgments

This thesis has been a four-year work at university of Rennes 1 in Lannion, France, supported by China Scholarship Council (CSC). I wish to express my deepest gratitude to my PhD supervisor, Prof. Daniel Chillet and associate professor Cédric Killian for their patient guidance and support throughout my PhD study. During discussions with them, they have always been able to explain to me clearly my questions and enlarge my vision. I really appreciate their time, advises, and effort to maintain my research on the right track. At the same time, I thank associate professor Sébastien Le Beux (Ecole Centrale de Lyon) for giving me a lot of constructive suggestions in this thesis.

I would like to thank all the members of the jury who evaluated this work: Olivier Romain (Professeur, Université de Cergy Pontoise), Cécile Belleudy (Maître de conférences, HDR, Université de Nice Sophia Antipolis), Laurent Vivien (Directeur de recherche, CNRS, Université Paris Sud), Jean Philippe Diguet (Directeur de Recherche CNRS, LabSticc Lorient), Camel Tanougast (Professeur, université de Lorraine).

I thank Prof. Olivier Sentieys for welcoming me as a PhD candidate to work in CAIRN team. I am grateful to his comments/suggestions during the research. I am very thankful to my colleagues, Van-Dung Pham, Ashraf El-Antably, Rengarajan Ragavan, Martha Johanna Sepúlveda, Christophe Huriaux who have been supportive and helpful during my Ph.D. work.

Finally, I would like to show my great appreciation to my family for their support in all these years. I would like thank my parents and my brother. I would not have started my PhD research in France without their support. I must acknowledge my husband, Aurélien Lebreton. I would not have finished this thesis without his understanding and encouragement. Specially, I really appreciate the birth of my dear daughter, Miya.

Her smile is the best gift for me everyday. My life is full of happiness with my lovely daughter.

Contents

Abreviations

AG	Architecture Graph
BER	Bit Error Ratio
CG	Conflict Graph
CMOS	Complementary Metal-Oxide-Semiconductor
FSR	Free Spectral Range
GA	Genetic Algorithm
IP	Intellectual Property
MPSoC	Multi-Processor System-on-Chip
MR	Microring Resonator
MWMR	Multiple Writer Multiple Reader
MWSR	Multiple Writer Single Reader
NN	Normalized Noise
NoC	Network-on-Chip
ONoC	Optical Network-on-Chip
PCM	Photonic Crystal Mirror

SNR	Signal to Noise
SoC	System-on-Chip
SOI	Si-on-Insulator
SUOR	Sectioned Undirectional Optical Ring
SWMR	Single Writer Multiple Reader
SWSR	Single Writer Single Reader
TbNoC	Tree-based Network-on-Chip
TG	Task Graph
TSV	Through-Silicon Via
VCD	Value Charge Dump
VCSEL	Vertical Cavity Surface Emitting Laser
WA	Wavelength Allocation
WDM	Wavelength Division Multiplexing
WiNoC	Wireless Network-on-Chip

List of Figures

List of Tables

Abstract

Evolution of Multi-Processor System-on-Chip (MPSoC) is moving towards the integration of hundreds of cores on a single chip. Designing an efficient interconnect for such complex architectures is challenging due to the ever growing data exchange between processors. In this context, silicon photonics is proposed as an emerging technology for future generation on-chip interconnects, providing several prospective advantages such as low transmission latency and high bandwidth. Optical Network-on-Chip (ONoC) are based on waveguides, carrying optical signals, and optical devices allowing to inject or drop the signals into this waveguide from an electrical interface. The waveguide can support multiple communications at the same time on different wavelengths by using Wavelength Division Multiplexing (WDM), providing a significant increase in bandwidth. However, simultaneous transmissions, on close adjacent wavelengths, may introduce inter-channel crosstalk noise through different optical switching elements within the network, which results in high Bit Error Ratio (BER). Besides, the higher the number of signals propagating simultaneously, the higher the crosstalk and the higher the laser output power needs. This leads to the following conflicting objectives: high-performance tends to rely on an exhaustive use of the wavelength while low-power communication prefers low bandwidth channel. Using an ONoC to execute a given application is thus a tedious task, especially if performance, power and BER objectives are likely to evolve with the execution context.

In this thesis, we assess communication performance and energy efficiency of ONoC from offline and online wavelength allocation respectively. Firstly, we propose an offline methodology which combines wavelength allocation and laser power scaling to find the best performance and energy consumption trade-off for ONoC-based MPSoC. The methodology is formed as a framework handling several inputs, such as architecture, application, mapping, and device parameters. It allows not only the exploration of

design space at both device level and system level, but also the generation of performance and energy efficiency trade-off based on a set of defined input parameters. The resulting Pareto solutions include: 1) low-power solutions, which tend to minimize the number of used wavelengths, and 2) high-performance solutions, for which multiple wavelengths are allocated to shorten the communication time. As an example, for a 107-task application running on 256 cores, the relative variation in the execution time and energy is 69% and 40% respectively. Finally, these solutions can be embedded in the controller and deployed at run-time according to execution requirement (e.g. high performance and low power). Secondly, in order to handle a dynamic workload and run-time deployment of tasks, dynamic policies to allocate the optical resources at run-time in a MPSoC architectures should be considered. We start by proposing a tree-based electrical set-up control network, named TbNoC, that centrally handle the requests to reduce channel allocation set-up latency. A minimization of crosstalk algorithm, which allocates distinct and separate wavelengths for optical signals, is employed to reduce the crosstalk, hence, the energy efficiency is improved. TbNoC is generic and can be combined to different ONoCs based on shared waveguide, such as Chameleon and SUOR. As a result, we explore the design space of TbNoC in terms of the number of tree levels and the number of wavelengths. To illustrate the SNR of TbNoC, we instantiate TbNoC to Chameleon ONoC, named Chameleon-TbNoC. As a result, we show that Chameleon-TbNoC improve a SNR of 1.6dB compared to SUOR under uniform traffic. In summarize, how to allocate efficiently wavelengths to improve energy efficiency and performance in ONoC are extensively discussed and handled in this thesis.

Keywords: Optical network-on-chip, wavelength allocation, energy efficiency, crosstalk noise, design exploration

Chapter 1

Résumé étendu

1.1 Contexte

Le système sur puce (SoC pour System on Chip) actuellement dévéloppés contiennent des milliards de transistors sur une seule puce de silicium mettant en œuvre diverses fonctions (par exemple, numérique, analogique et mixte, etc.). De nos jours, les SoC sont déployés dans de nombreux appareils électroniques numériques tels que les téléphones mobiles, les consoles de jeux vidéo, les lecteurs MP3, etc. La conception de ces SoCs devient de plus en plus complexes en raison de la complexité des applications. Au fur et à mesure du développement de la technologie des semi-conducteurs, le nombre de cœurs intégrés sur une seule puce augmente régulièrement. C'est anisi que de nos jours, des centaines de cœurs peuvent exécuter de nombreux processus sur une seule puce, appelée Multi-Processor SoC (MPSoC). À mesure que le nombre de cœurs augmente, cela entraîne de plus en plus de transferts entre eux, et ces transferts augmentent considérablement l'énergie nécessaire pour déplacer les données d'un cœur à l'autre. Le nombre de transfers augments considérablement. En conséquence, la bande passante devrait atteindre les dizaines et centaines de térabits par seconde [12]. Malheuresement, ces transferts engendrent une augmentation importante de l'énergie. En outre, lorsque la communication globale sur puce augmente, des conflits de ressources peuvent entraîner une latence de réseau élevée. Par conséquent, une interconnexion efficace est demandée pour assurer une communication efficace et fiable.

La photonique sur silicium a le potentiel pour surmonter cette augmentation du flux de communication en remplaçant les fils électriques par un guide d'ondes optique. Les avantages de l'interconnexion optique sont détaillés ci-dessous. La maturité de la fabrication photonique sur silicium offre la possibilité d'une intégration à grande échelle sur puce. Par example, STMicroelectronics a développé une plate-forme de technologie photonique silicium à faible coût [28]. Intel a fait la démonstration d'une liaison photonique sur silicium de 4 canaux à 12.5Gbps avec un laser hybride intégré au silicium [29]. IBM a fait la démonstration de switch photoniques rapides en silicium à base de Mach-Zehnder [30]. Par ailleurs, le délai de propagation dans le guide d'onde est simplement le temps de vol à la vitesse de la lumière. Les routeurs électriques doivent switcher chaque bit de donnés transmis, entraînant une augmentation de la puissance dynamique en fonction du débit et de la distance de communication, tandis que les switches photoniques s'allument et s'éteignent une fois par message, sans buffering ni traitement, ce qui améliore le rendement énergétique (référé comme la transparence de débit binaire dans les interconnections optiques) [2]. Un autre avantage important est l'utilisation de plusieurs longueurs d'onde dans un guide d'ondes partagé qui permet d'offrir une bande passante élevée. La modulation sur une seule longueur d'onde de 10Gb/s a été démontrée dans [32] et 64 longueurs d'onde par guide d'onde sont présentées dans [9], donc la bande passante totale de 640Gb/s peut être atteinte avec un seul guide tandis que dans un NoC électrique, une telle augmentation ne peut se faire que par augmentation du nombre de fils pour le support de communication.

Dans cette thèse, nous considérons une architecture intégré 3D comme une plate-forme pour prendre en charge la technologie optique NoC: 3D ONoC. Il combine les avantages de la photonique sur silicium et de l'empilement 3D pour surmonter les limitations traditionnelles du réseau électrique. L'interconnexion photonique fournit une bande passante élevée sur la puce et une faible latence tandis que l'empilement 3D peut réduire la distance d'interconnexion. Une interconnection optique est composée d'un émetteur (source laser et modulateur), d'un support de transmission (guide d'onde) et d'un récepteur (photodétecteur). En général, une source laser émet une lumière optique, puis un modulateur convertit les informations numériques du domaine électrique en domaine optique en fonction des bits de données ("1" ou "0"). Après modulation, le signal optique se propage dans le guide d'onde de la source au récepteur où le signal optique est reconverti dans le domaine électrique. Alors que ONoC peut offrir des opportunités intéressantes, plusieurs défis doivent être adressés pour exploiter efficacement ses capacités.

1.2 Challenges autour des ONoC

L'interconnexion photonique sur silicium peut offrir des avantages tels qu'une faible latence de transmission et une bande passante élevée. Cependant, les dispositifs photoniques introduisent de nouveaux effets électromagnétiques (perte d'insertion et crosstalk), sans équivalent électronique, qui ont un impact important sur l'efficacité énergétique des ONoC. En outre, les lasers sont la principale source de consommation d'énergie des ONoC. Par conséquent, dans cette thèse, nous nous concentrons sur l'amélioration de l'efficacité énergétique de l'interconnexion photonique en adaptant la puissance du laser, en tenant compte du crosstalk et du taux d'erreur binaire (BER) cible pour la communication. En outre, le compromis entre l'efficacité énergétique et la performance de communication est également étudié.

1.2.1 Crosstalk

Le crosstalk est introduit par l'interférence de différents signaux dégradant la qualité de la transmission. Les solutions standards au niveau de la puce sont proposées pour réduire le crosstalk dans l'interconnexion photonique. Par exemple, Sanchis *et al.* a proposé de minimiser le crosstalk dans les guides d'ondes SOI en choisissant l'angle de croisement optimal. Les résultats expérimentaux ont montré que l'angle de croisement optimal n'est pas de 90 degrés mais de 60 ou 120 degrés [64]. Jayatilleka *et al.* ont étudié le crosstalk du premier et deuxième ordre introduit par MR. Ils démontrent que le crosstalk peut être réduit en augmentant l'ordre du MR [65].

Le bruit de crosstalk au niveau de la puce peut sembler négligeable, mais le crosstalk accumulée au niveau du réseau peut gravement affecter les performances et le rapport signal sur bruit (SNR) du réseau. Certaines recherches consacrées à l'exploration du crosstalk au niveau du réseau ont été effectuées. Xie *et al.* a étudié le crosstalk, SNR et le taux d'erreur binaire (BER) du réseau mesh. Un routeur optique Crux est proposé pour minimiser la perte d'insertion et le crosstalk [66]. Nikdast *et al.* ont analysé le crosstalk et SNR dans folded-torus-based [67] and fat-tree-based ONoC [68] reseau. Les études dans la littérature ci-dessus se concentrent généralement sur le pire cas en termes de demandes sur la technologie. Fusella *et al.* a proposé un outil de mappage pour atténuer le crosstalk en réduisant le nombre de communications partageant un guide d'ondes [69]. Ces modèles mentionnés ci-dessus utilisent seulement une approche à longueur d'onde unique et le crosstalk provient d'une seule longueur d'onde. Cependant, le SNR est inversement proportionnel au nombre de longueurs d'onde utilisées dans le réseau et le problème de crosstalk est plus critique dans le cadre de l'utilisation du WDM ONoC. Duong *et al.* a étudié le crosstalk dans un ONoC à anneau de multiplexage DWDM. Ils ont démontré qu'un facteur de qualité plus élevé peut réduire le crosstalk, améliorant ainsi l'efficacité énergétique [70]. Chan *et al.* a proposé un simulateur de réseau optique, appelé PhoenixSim [48], pour évaluer davantage l'espace de conception du réseau photonique. Ils ont présenté une méthodologie pour explorer conjointement les conceptions au niveau physique et au niveau du système des topologies de réseau [38]. Ces propositions étudient la consommation d'énergie pour plusieurs topologies existantes, mais n'utilisent pas cette information pour réduire la consommation d'énergie du réseau.

La puissance de crosstalk reçue par le photodétecteur dépend de la fréquence de coupure du MR: plus les autres signaux sont proches de la longueur d'onde de résonance, plus élevé sera crosstak. Chittamuru *et al.* proposent une technique de réduction du crosstalk pour augmenter l'espacement des canaux entre les longueurs d'onde adjacentes dans un réseau DWDM [72]. Cependant, cette technique d'espacement des longueurs d'onde conduit à la réduction du degré du nombre de canaux DWDM dans le guide d'onde, à son tour, réduisant la bande passante du canal et le débit du réseau.

1.2.2 Gestion de puissance de Laser

Gérer un réseau ONoC pour exécuter une application est une tâche fastidieuse, en particulier lorsque la consommation d'énergie est étroitement liée aux objectifs du BER. La consommation d'énergie des lasers est une partie importante de la consommation d'électricité statique ONoC. Pour améliorer l'efficacité énergétique, certaines approches présentées dans la littérature proposent de régler la puissance du laser. Zhou *et al.* a proposé une stratégie de prédiction basée sur l'utilisation du réseau pour moduler la puissance du laser: la puissance du laser est ajustée en fonction des prédictions basées sur l'activité précédente [73]. Ils ont conçu un schéma efficace avec un prédicteur de traffic de bande passante pour augmenter l'utilisation du canal et réduire la puissance optique statique dans le laser externe. Les lasers sur puce sont plus économes en électricité statique que les sources laser hors puce. Étant donné que le laser sur puce peut être mis hors tension lorsqu'il n'y a pas de transfert de données, cela réduit considérablement la consommation d'énergie si la charge de l'application n'est pas élevée. Dans [74], une boucle de contre-réaction configure la puissance du laser dans un mode linéaire en utilisant un DAC à 6 bits. Cependant, cette solution souffre d'une mauvaise évolutivité en raison de la nécessité d'estimer le BER du côté du récepteur et du contrôle de retour. De plus, il n'y a pas d'analyse de la pénalité induite par l'estimation du BER. En général, la dissipation de puissance du laser dépend de la bande passante du canal de données et des pertes photoniques le long du chemin optique. Afin de réduire avec précision la dissipation de puissance, Chen *et al.* a proposé une stratégie on-off pour maximiser l'efficacité énergétique [75]. Les liaisons optiques peuvent accéder à la mémoire (caches L2) et peuvent être activées / désactivées au moment de l'exécution en fonction de la taille de l'application. Selon les résultats de simulation, la consommation moyenne d'énergie du laser sur puce est réduite de 23,8% en utilisant une stratégie de contrôle on-off par rapport à la stratégie qui maintient toujours les lasers allumés. Cette méthode peut être efficacement combinée à la stratégie de partage et de placement laser qu'ils proposent pour améliorer l'efficacité énergétique de l'ONoC [76]. Dans [77], un CMOS driver est conçu pour configurer l'alimentation laser en mode OFF, veille, moyenne et pleine puissance. La puissance du laser est commutée entre les quatre modes pour fournir une puissance optique juste suffisante pour la bande passante de communication souhaitée et pour le BER cible. Plus la bande passante de communication est élevée, plus la puissance du laser est élevée. Ils considèrent également une pénalité de temps et de puissance pour le mode OFF et le mode veille respectivement.

1.3 Contributions

Les travaux publiés dans la littérature abordent plusieurs aspects intéressants de l'ONoC, mais aucun d'entre eux ne propose une approche globale pour assurer l'allocation de longueur d'onde, la gestion laser et la gestion en ligne hors ligne. Compte tenu de cet état de l'art, cette thèse propose plusieurs contributions.

La contribution principale est d'allouer une ou plusieurs longueurs d'onde pour chaque communication, en considérant conjointement la position la plus appropriée de chaque longueur d'onde réservée pour améliorer l'efficacité énergétique du réseau. Le problème d'allocation de longueur d'onde est résolu par une approche hors ligne et en ligne respectivement.

Dans l'approche hors ligne, nous combinons l'allocation de longueur d'onde et l'échelle de puissance laser pour trouver le compromis entre la performance et la consommation d'énergie. La méthodologie est formalisée sous forme d'un framework prenant compte de plusieurs entrées, telles que l'application, l'architecture et les paramètres de la puce. D'une part, l'exploration de l'espace de conception au niveau du dispositif et du système est autorisée; d'un autre côté, un ensemble de compromis entre performance et efficacité énergétique peut être généré une fois les paramètres d'entrée définis. Les solutions Pareto qui en résultent comprennent 1) des solutions de faible puissance, qui tendent à minimiser le nombre de longueurs d'onde utilisées, et 2) des solutions hautes performances, pour lesquelles plusieurs longueurs d'onde sont allouées pour raccourcir le temps de communication. Enfin, ces solutions peuvent être intégrées dans un contrôleur et déployées lors de l'exécution en fonction des exigences d'exécution (par exemple haute performance et faible consommation).

Cependant, l'approche hors ligne est limitée à une application prédéfinie, c'est-à-dire qu'elle n'est pas capable de gérer une charge de travail dynamique et un déploiement de tâches au moment de l'exécution. Dans ce cas, des politiques dynamiques pour allouer les ressources optiques à l'exécution dans une architecture MPSoC doivent être considérées pour cet objectif. Nous proposons un réseau de contrôle de configuration électrique arborescente, TbNoC, pour la gestion en ligne de l'ONoC. Nous partitionnons le réseau de contrôle électrique global en plusieurs agents hiérarchiques centralisés et centralisons la ressource d'arbitrage dans le but de réduire la latence du chemin d'installation. Ensuite, un algorithme de minimisation du crosstalk est utilisé pour séparer les longueurs d'onde des signaux optiques afin de réduire le crosstalk, par conséquent, le rendement énergétique est optimisé. Enfin, le mécanisme de ressources entièrement partagées est plus flexible que les règles d'accès limité mises en œuvre dans Corona et SUOR, mais il en résulte une plus grande complexité du réseau. Par conséquent, un compromis entre la flexibilité et la complexité du réseau doit être envisagé par le concepteur.

Dans notre thèse, l'allocation de longueur d'onde peut être spécifiée au moment de la compilation ou réalisée au moment de l'exécution. Dans le premier cas, les demandes de communication sont connues à l'avance, puis une simulation hors ligne supplémentaire est nécessaire pour sélectionner les longueurs d'onde appropriées. Dans ce dernier cas, une séquence de demandes de communication est reçue au moment de l'exécution et un chemin optique est établi en réponse à chaque demande. Chaque approche est adaptée à différents scénarios d'application et les deux méthodes pourraient être utilisées conjointement dans l'ONoC.

Chapter 1

Introduction

1.1 Evolution of System-on-Chip Architecture

System-on-Chip (SoC) contains billions of transistors on an unique chip implementing various functions (e.g. digital, analog, and mixed-signal etc). Nowadays, SoCs can be found in many digital electronic devices such as mobile phones, video game consoles, and MP3 players etc. These SoC designs become more complex due to the increasing complexity of applications. As semiconductor technology develops, the number of cores integrated onto a single chip is expected to climb steadily as shown in figure 1.1. Nowadays, hundreds of cores can run multiple processes on a single chip, called Multi-Processor SoC (MPSoC). Some recent proposed architectures for embedded systems, such as Kalray's MPPA (targeting 1024 cores, current version is 256 cores), can be found in a multitude of applications including medical devices, aeronautics, automobiles, and high-performance data centers [11]. As the number of cores increases, it leads to more and more transfers between them, and these transfers drastically increase the energy to move data from one core to another. As a consequence, the bandwidth is expected to reach the tens and hundreds of terabits per second for future on-chip communication [12]. Furthermore, when global on-chip communication becomes severe, high resource contention may cause high network latency. As a result, an efficient interconnect is expected to ensure effective and reliable communication.

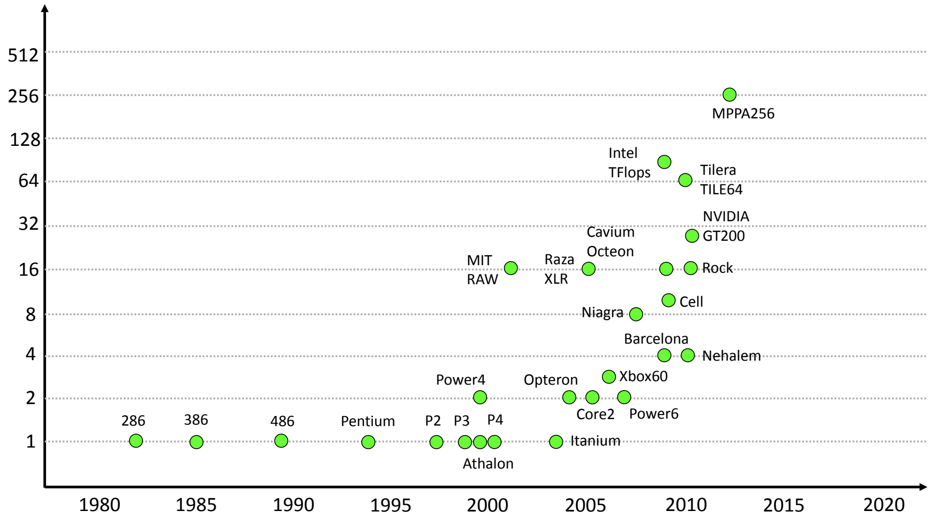


Figure 1.1: Evolution of system-on-chip architecture [1].

1.1.1 Point-to-Point and Bus

Point-to-Point architecture is implemented by connecting all Intellectual Property (IP) cores with a dedicated communication channel. However, the Point-to-Point architecture lacks scalability due to high design effort, cost, and complexity [13]. Connecting all the IPs on a single bus is another solution for interconnect. In this case, an IP core broadcasts data on this bus and all other IPs can read data from it. Since the single bus is shared by all the cores, only one core is allowed to use it at a time. Hence, the arbitration to decide which core can transmit data is indispensable to avoid contentions. However, the inefficiency of running parallel applications is evident due to arbitration among possible traffic flows. Besides, bandwidth limit of a bus leads to growing wiring delay [14].

1.1.2 Network-on-Chip

To alleviate complex communication problems that arise as the number of on-chip components increases, Network-on-Chip (NoC) has emerged. NoCs support data transmission as packets via a network that consists of routers, network interfaces, and links. Another main parameter affecting the overall network performance is the topology which specifies the physical organization of interconnection network, such as simple ring, two-dimensional mesh, fat tree, etc. NoC is accepted as the preferable communication backbone enabling the integration of large numbers of embedded cores on a single chip for MPSoC. Wolkotte *et al.* demonstrated a lower energy consumption per bit for NoC architecture compared to a traditional bus architecture, i.e 5pj/bit for 25 nodes NoC in contrast with 32pj/bit for 25 nodes bus [15]. However, bandwidth density of NoC is largely limited by the pitch of electrical wires. Moreover, as the system size grows, if a given packet traverses a large number of hops to reach its destination, NoC latency is long [16]. Finally, a remarkable increase in power dissipation of electrical links is demonstrated since it tends to scale with throughput performance. As reported in [17], NoC switching activity uses over 50% of the interconnect power consumption. Moreover, 80% of chip power is consumed by on-chip interconnects [18]. This is mainly due to the projected increasing in global wire power dissipation and decreasing gate power consumption. Consequently, any future interconnect technology should be evaluated on three key metrics: bandwidth density, latency, and energy efficiency.

1.2 Emerging Interconnect Paradigms

To improve performance and energy efficiency of on-chip interconnect, three major emerging technologies are proposed: 3D technology, wireless NoC, and silicon photonics.

1.2.1 3D Technology

3D packaging technology shifts the manufacture from horizontal to vertical dimension. It allows multi-dies stacked above each other and vertically interconnected using Through-Silicon Vias (TSVs) [19] (or silicon interposer [20], monolithic 3D [21]) on the same chip. Each die may integrate different technologies, such as photonic, logic, DRAM, non-volatile memory, and so on [9]. It overcomes limitation of scaling on 2D-NoC, since a large die can be partitioned into small segments and stacked vertically on top of each other. The shift from 2D-NoC to 3D-NoC improves electrical characteristics (such as latency and power) by reducing wire connection lengths between devices [22]. A decrease of 40% number of hops is demonstrated in [16], then a decrease of 40% in zero-load latency and 62 % in power consumption is reached with 3D-NoC compared to traditional 2D-NoC for a network size of 128 nodes [23]. However, this electrical approach does not address the difficulties such as voltage isolation, wave reflection, impedance matching, PIN inductance, timing accuracy, and overall ease of design [24]. Therefore, changing the physical means of interconnection, like radio frequency and nanophotonics, could be an attractive alternative to traditional metallic wires for future MPSoC.

1.2.2 Wireless Network-on-Chip

The major challenges in wire-based on-chip communication are high latency and power consumption of their multi-hop links [2]. This problem will be further exacerbated with an increase in the number of cores on a chip. To alleviate the problem of multi-hop communication links, the integration of a high-bandwidth single-hop long-range wireless link is proposed. The wireless interconnect provides benefits including: (1) high energy efficiency for single-hop long-rang communication, (2) reduced complexity compared to traditional wire-based system, and (3) compatible with Complementary Metal-Oxide-Semiconductor (CMOS) [25]. The Wireless NoC (WiNoC) is based on on-chip antennas, transceivers, and routers [26]. The feasibility of integrating compact antennas is demonstrated in [27], which allows to operate in the MilliMeter (mm)-wave range of a few tens to one hundred GHz for intra- and inter-chip communication. However, the antenna size in mm-wave range is still a limitation for the implementation of on-chip communication (i.e 1-2mm [2]). Having wireless nodes spread all over the chip is not that economical due to the antenna and transceiver footprint. In this context, a hybrid (wireless/wired) and hierarchical NoC architecture WiNoC is proposed as depicted in figure 1.2, where wireless is used to connect subnets (a group of cores) and each subnet is connected using wired connected network. However, the frequency spectrum cannot be reused due to the nature of frequency sharing among different subnets [25].

1.2.3 Silicon Photonic

Silicon photonics has the potential to overcome the communication bottleneck by replacing electrical wires with optical waveguide. The advantages and drawbacks of photonic interconnect are detailed in the following.

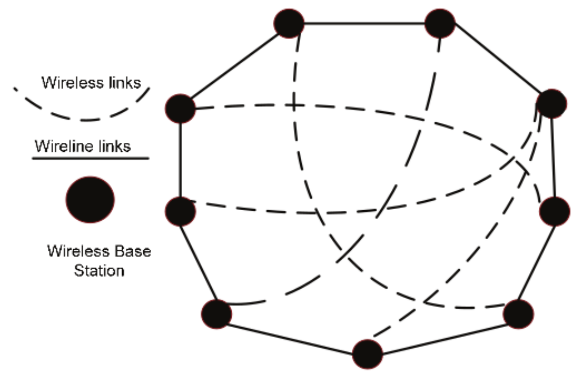


Figure 1.2: Illustration of WiNoC [2].

Advantages: the maturity of silicon photonic manufacturing offers the possibility of large-scale integration on chip. STMicroelectronics developed a low cost silicon photonic technology platform [28]. Intel demonstrated a 4-channel × 12.5Gbps silicon photonic link with integrated hybrid silicon laser [29]. IBM demonstrated a fast Mach-Zehnder-based silicon photonic switches [30]. Besides, the propagation delay on waveguide is simply the time of flight at the speed of light [31]. Electric routers must switch with every bit of transmitted data, resulting in a dynamic power dissipation scaling with the bit rate, while photonic switches turn on and off once per message, without buffering and processing, resulting in an improved energy efficiency independent from the bit rate (referred as bit-rate transparency in photonic community) [2]. Another significant advantage is the use of multiple wavelengths in a shared waveguide that offers high bandwidth density. High-speed 10Gb/s single wavelength modulation has been demonstrated in [32] and 64 wavelengths per waveguide are anticipated in [9], thus the aggregate bandwidth of 640Gb/s can be reached using only one waveguide, while electrical NoC requires to increase the number of wires to increase bandwidth.

Drawbacks: due to technology limitations, silicon photonic still encounters challenges on design side, such as electrical/optical codesign and cosimulation. Two other key technical challenges facing photonic interconnect are the sensitivity of photonic devices to process and temperature variations. The existence of inherent spatial variations in CMOS materials and fabrication processes degrade more greatly silicon photonic devices than electronic devices [33]. Another major limitation is that the optical signal can not be buffered and processed directly. Therefore, Optical to Electrical (OE) conversion is required in the photonic interconnect, in turn, an extra part of energy is consumed. Finally, photonic introduces new electro-magnetic effects (insertion loss and crosstalk) that greatly impact the photonic interconnect [31].

1.3 Challenges and Contributions

In this thesis, we consider 3D integrated circuit as a platform to support mixed-technology electronic-controlled Optical NoC: 3D ONoC. It combines the benefits of silicon photonic and 3D stacking to overcome traditional electrical network limitations. Photonic interconnect delivers high on-chip bandwidth and low latency while 3D stacking can reduce the interconnect distance. Figure 1.3 shows a ring-based 3D ONoC architecture composed of two layers interconnected by TSV: i) on the bottom, an electrical layer implementing processing cores and ii) on the top, an optical layer with optical routers and waveguides. While this architecture model can offer interesting opportunities, several challenges must be addressed to efficiently exploit its abilities.

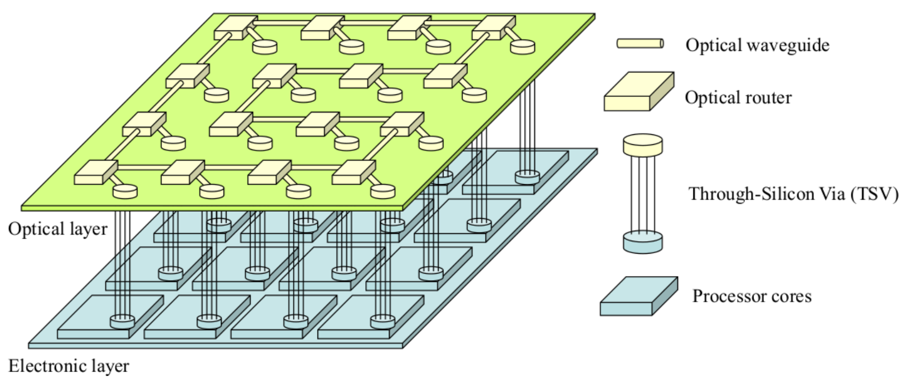


Figure 1.3: Three-dimensional architecture of a ring-based ONoC [3].

Challenges: one main concern of optical network is the limited on-chip power budget for current MPSoC. The basic optical devices employed to construct ONoC are not perfect. When an optical signal passes through optical devices, a small amount of optical power is filtered into them, called insertion loss. The total loss in any optical link is calculated as the sum of losses of all the physical components which affects optical signal. In addition, optical signals of various wavelengths can interfere with each other through optical switching elements, introducing crosstalk noise further aggravating Signal to Noise Ratio (SNR), which can be compensated by increasing the laser power. This leads to increase power consumption. The higher the number of wavelengths per channel, the higher the crosstalk and the higher the laser output power needed to satisfy specific Bit-Error-Ratio (BER) requirement. To minimize overall power consumption, it is important to use low laser power for low bandwidth channel.

Another concern of optical network is lack of data buffering structures in optical domain. Usually, optical switch turns on/off once per message, a preliminary wavelength allocation is thus needed before data transfer, then the optical data can be transmitted in an end-to-end manner. Furthermore, the use of multiple wavelengths in a shared waveguide can offer high bandwidth channels to reduce transmission time. However, the higher the number of reserved wavelengths for ongoing communications, the smaller the number of free wavelengths for future communications.

In conclusion, high-performance tends to rely on an exhaustive use of the wavelengths while low-power communication prefers low bandwidth channel. Thus the way to manage the energy and performance trade-off in a shared waveguide is critical to fully exploit the benefits of photonic interconnects.

Contributions: this thesis assesses the performance and energy efficiency of ONoC from offline and online wavelength allocation respectively. In the former case, the communication needs are known in advance; while in the latter case, the optical paths are set-up on demand at run-time. Firstly, we propose an offline methodology which combines wavelength allocation and laser power scaling to find the best trade-off for performance and energy consumption. The methodology is formed as a framework to generate ONoC configurations. Then, the trade-off configurations can be embedded in the controller and deployed at run-time according to execution requirement. Secondly, to address online wavelength allocation problem, a tree-based electrical control network is designed to centrally handle set-up requests at run-time for purpose of reducing set-up latency. In particular, minimization of crosstalk algorithm is employed to improve SNR of system, which allows to reduce the optical power emitted by lasers and hence their power consumption. In summarize, how to allocate efficiently wavelengths to improve energy efficiency and performance in ONoC are extensively discussed and handled in this thesis.

1.4 Thesis Outline

The thesis is organized in the following five chapters:

In chapter 2, we review related preliminary knowledge of on-chip silicon photonic interconnects. Firstly, general optical link and optical devices employed on silicon photonic interconnect are introduced. Then, shared channel designs and existed ONoC architectures are outlined. Finally, discussion of our propositions are presented.

In chapter 3, offline wavelength allocation and laser power tuning are combined to explore the performance and energy consumption trade-off for ONoC-based MPSoC. This chapter begins with key concepts of wavelength allocation and laser power configuration, then a multi-objective formalization is presented. From this formalization, we propose a framework, including problem descriptions, which produces a set of ONoC configurations according to user specification.

In chapter 4, as a framework is proposed to generate performance and energy efficiency trade-off for ONoC-based MPSoC, architectural exploration on a given application at both device level and system level is given at first. The impact of microring resonator characteristic parameters, number of laser power levels, number of wavelengths, and task mapping are investigated. Then, the scalability of the methodology using ONoC architectures with 64, 128, and 256 cores is studied.

In chapter 5, a tree-based electrical control network for online management of ONoC is introduced. This network, named TbNoC, reduces optical channel set-up delay and improves SNR performance. Firstly, key concepts of design including partition principle and communication protocol are introduced. Then implementation of architecture is given. Finally, optical channel set-up delay and SNR performance of proposed architecture are evaluated.

Finally, conclusions are drawn and perspectives of this thesis are presented in Chapter 6.

Chapter 2

Related Work

As core number increases to hundreds in future MPSoC, high performance and energy efficient interconnects are required. Silicon photonic links have been considered as one of potential solution to address challenges associated with performance scaling of the computing system from MPSoC. In this chapter, section 2.1 introduces photonic interconnect basics, including laser, waveguide, Microring Resonator (MR), and photodetector. Then we detail different shared channel designs in section 2.2, and WDM-based ONoC architectures in section 2.3. Section 2.4 presents the BER estimation model while sections 2.5 and 2.6 discuss the position of our contribution compared to the published work. Finally, section 2.7 concludes this chapter.

2.1 Photonic Interconnect Basics

The specific implementations of photonic interconnects can vary substantially, but from a high level view, a general optical link is basically composed of: a transmitter (laser source and modulator), a transmission medium (waveguide), and a receiver (photodetector). As shown in figure 2.1, a laser source emits optical light, then a modulator converts digital informations from electrical domain to optical domain according to data bits ("1" or "0"). After modulation, the optical signal propagates into the waveguide from source to receiver where the optical signal is converted back into electrical domain. In this section, optical devices employed to construct photonic interconnect, such as laser, waveguide, microring resonator, and photodetector, are detailed in the following subsections.

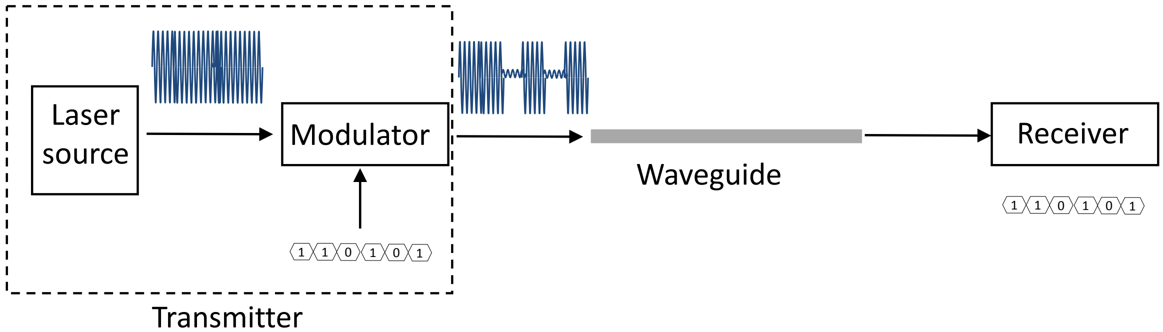


Figure 2.1: Generic communication in photonic interconnect.

2.1.1 Laser

Two types of laser sources are classified according to their location: off-chip and on-chip. There is debate whether the optical source should remain off-chip or be integrated on the processor chip. Figure 2.2 illustrates three point-to-point photonic links. In link (a), an off-chip laser source generates optical light on the wavelength λ_x coupled into waveguide with an optical fiber. After being modulated, the optical signal (mark ②) propagates into the waveguide from source to destination. However, the optical signal suffers from the loss from couple fibre-waveguide and from optical devices along the optical path. Then, at destination side, ON-state filter specific to λ_x drops optical power (mark ③) into photodetector who converts the optical signal back into electrical domain (mark ④). This solution, however, causes energy waste when the off-chip laser is not used and cannot be turned off. In link (b), an on-chip laser is integrated within the chip and emits a constant light at λ_x into the waveguide. Then this optical signal is also modulated by a modulator, as in link (a), and sent to the receiver. Differently, link (c) employs an on-chip laser using direct modulation. In this way, the modulated optical signal is directly emitted by the on-chip laser. Besides, this on-chip laser can be powered off when there is no data transfer. In our thesis, we consider the use of on-chip laser with direct modulation due to its high flexibility. Vertical Cavity Surface Emitting Lasers (VCSELs) are cost-effective solutions and can be integrated due to their low threshold and driving current [34]. As detailed in [35], on-chip lasers operate in three modes: 1) OFF, 2) linear, and 3) saturation. The current for which the gain satisfies the lasing condition is the threshold current of the laser. When driving current is below the threshold, no optical power is emitted. When driving current is higher than

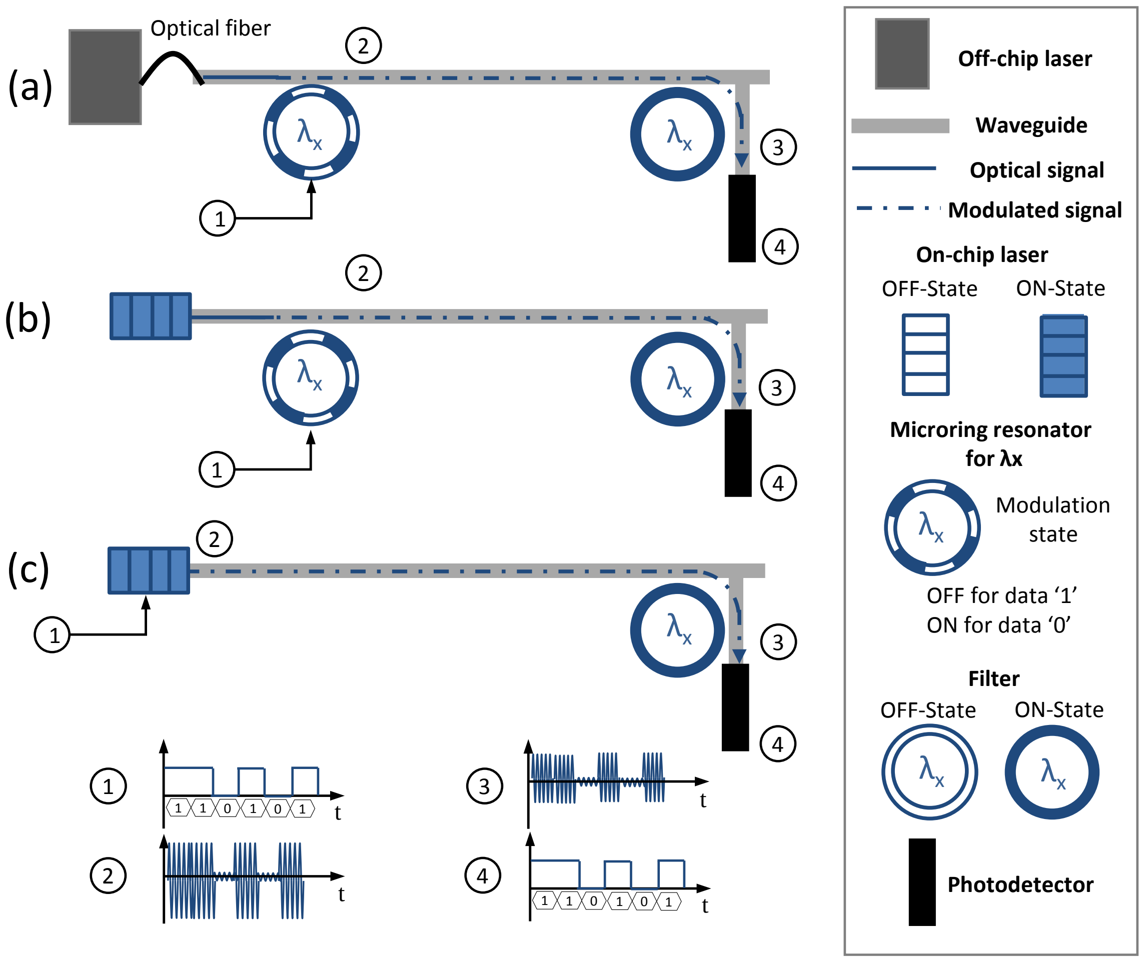


Figure 2.2: Silicon photonic interconnects with (a) external laser source using indirect modulation (b) on-chip laser using indirect modulation (c) on-chip laser using direct modulation. (① represents the electrical signal before modulation, ② represents the optical signal after modulation, ③ represents the optical signal before photodetection, and ④ represents the electrical signal after photodetection.

the threshold, the output power increases linearly with driving current until it arrives at its saturation point. Furthermore, VCSELs based on double set of Si/SiO_2 Photonic Crystal Mirror (PCM) are CMOS-compatible as shown in figure 2.3 [4]. Commercial VCSELs, when forward biased at a voltage well above 1.5V, can emit optical power of the order of a few mW around 850nm, with an efficiency of 40% [36]. The output power of on-chip laser can be dramatically controlled through run-time laser power management technique. However, one disadvantage of on-chip laser is the low laser efficiency, which is inversely proportional to the temperature.

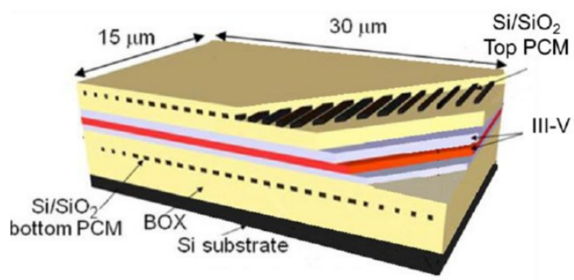


Figure 2.3: Illustration of PCM-VCSEL [4].

2.1.2 Waveguide

Waveguide is used as optical wires to confine and guide light. The most straightforward way to construct a waveguide is to use high refractive index as core and low refractive index as cladding, then the waveguide keeps a tight confinement of the light for transmission due to high refractive index contrast as shown in figure 2.4 [5]. All the various optical components, sources, and destinations are linked by the waveguide. Optical signal travelling through the waveguide suffers from propagation loss due to light scattering from the etched sidewalls. Increasing waveguide width and decreasing etched depth can reduce the waveguide propagation loss [37]. Dong *et al.* have demonstrated low loss shallow-ridge silicon waveguides with an average propagation loss of 0.274dB/cm [37]. In addition, waveguide bends have an important role in the creation of switches and topologies. However, it contributes to additional loss for the optical signal, which we refer as bending loss [38]. Tu *et al.* demonstrated that the bending loss per 180^o turn is 0.05dB for a bending radius of 5μm and 0.008dB for a bending radius of 10μm [39]. Finally, waveguide crossing occurs when two waveguides intersect to build different on-chip topologies [38]. Generally, electrical interconnects do not allow crossings of two wires due to short circuit effect. Waveguide crossing enables the creation of a large number of different topologies, however, the introduced insertion loss and crosstalk degrade the system performance. Zhang *et al.* demonstrated a waveguide crossing for submicron silicon waveguide with average insertion loss of 0.18dB $\pm$ 0.03dB and crosstalk of -41dB $\pm$ 2dB [40].

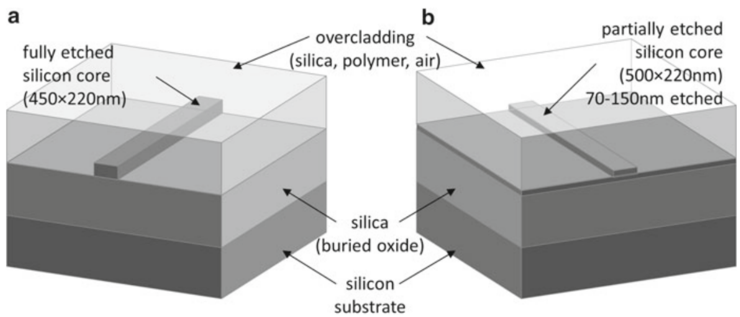


Figure 2.4: High-contrast silicon waveguide geometries. (a) "Photonic wire" strip waveguide, (b) rib waveguide [5].

2.1.3 Microring Resonator

Microring resonator (MR) is intensively employed to construct basic optical elements, such as modulators and filters, which will be detailed in the next sub-sections. A MR is characterized by its resonant wavelength. When the wavelength of an optical signal matches the resonant wavelength of a MR, the signal is coupled into the MR; otherwise, it passes through the MR.

2.1.3.1 Physical Characteristics

Fig. 2.5 depicts a typical MR with a desired radius R placed between two parallel waveguides.

A MR is characterized by its transfer coefficients k_e, k_d, k_p, and by its resonant wavelength λ_{MR}. For optical signals carried on wavelength λ_i, the power transferred to the drop port can be expressed as Eq.(2.1) [41]

$$\frac{P_{drop}^{\lambda_i}}{P_{in}^{\lambda_i}} = \left(\frac{2k_e k_d}{k_e^2 + k_d^2 + k_p^2}\right)^2 \left(\frac{\delta^2}{(\lambda_i - \lambda_{MR})^2 + \delta^2}\right) \tag{2.1}$$

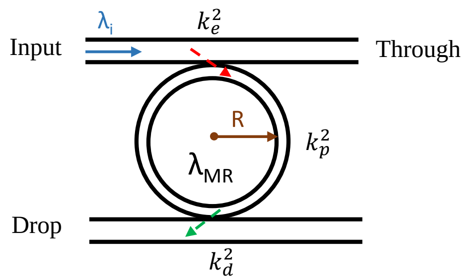


Figure 2.5: Theoretical representation of a microring resonator.

Where 2δ is the -3dB bandwidth of the MR that can be determined from the resonant wavelength and the quality factor (Q) of MR as expressed in Eq. (2.2).

$$2\delta = \frac{\lambda_{MR}}{Q} \tag{2.2}$$

In Fig. 2.5, k_e^2 is the fraction of the input optical power coupled into the MR and k_d^2 corresponds to the fraction of the MR optical power coupled into the drop port. Moreover, k_p^2 is the fraction of the intrinsic power losses per round trip in the MR. Without loss of generality, we suppose that $k_d^2 + k_e^2 >> k_p^2$ and $k_d \leq k_e$. As a result, the coefficient $\left(\frac{2k_e k_d}{k_e^2+k_d^2+k_p^2}\right)^2$ in Eq.(2.1) can be approximated to 1. Hence, we can simplify transmission function on drop and through ports (noted Φ^d and Φ^t) as below:

$$\Phi^d(\lambda_i, \lambda_{MR}) = \frac{\delta^2}{(\lambda_i - \lambda_{MR})^2 + \delta^2} \tag{2.3}$$

$$\Phi^t(\lambda_i, \lambda_{MR}) = 1 - \Phi^d(\lambda_i, \lambda_{MR}) \tag{2.4}$$

The resonant wavelength of a MR can be shifted by applying electro-optic [42] and thermo-optic [43] mechanism, leading to ON and OFF state of MR as shown in figure 2.6. Another parameter characterizing a MR is the Free Spectral Range (FSR) corresponding to the spacing between two successive resonances in the optical spectrum.

When an incoming signal centered at λ_{MR} reaches the input port of an ON-state MR characterized by its resonance wavelength λ_{MR}, this signal is redirected to the drop port (shown in figure 2.6 (a)), but a portion of power is leaked into through port (shown in figure 2.6 (b)).

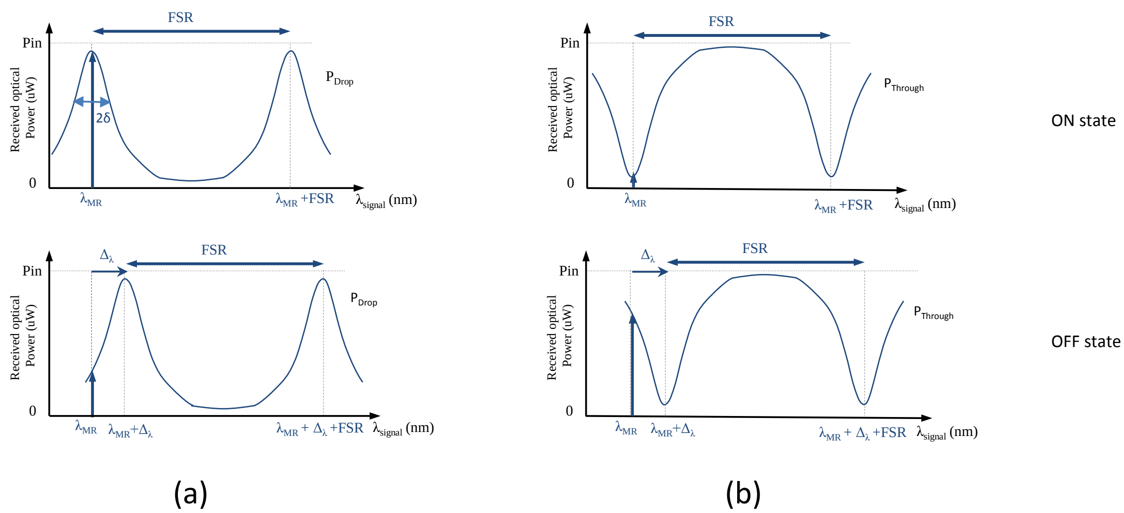


Figure 2.6: Signal transmission on (a) drop port and (b) trough port of ON and OFF state MR.

Differently, when this signal reaches the input port of an OFF-state MR, where its resonant wavelength is shifted at $\lambda_{MR} + \Delta\lambda$, a portion of power is leaked into drop port and the optical signal continues propagating to the through port.

2.1.3.2 Functional View

Microring Resonator as Modulators: a MR is capable of modulating specific wavelengths in the optical spectrum. MRs are always placed next to a waveguide, then they can couple the light into the ring and dissipate it. By configuring ON and OFF-resonance state, logical 0 and 1 are encoded into the optical light. To transmit '1', the wavelength is not modified when passing through the MR; to transmit '0', the wavelength is extracted by the MR and dissipate. 12.5Gbit/s modulation using a silicon MR modulator is demonstrated in [42]. Besides, MR modulators have the advantages of low power consumption and high efficiency. Zheng *et al.* demonstrated that the energy efficiency of cascaded tunable MR modulators is nearly 100fj/bit excluding laser power and tuning power [44].

Microring Resonator as Filter: MR is used as a filter to selectively extract a single wavelength from the input waveguide. Light coupled into the ring will not dissipate like the modulator, but will be redirected to another waveguide. Filtering a wavelength can be achieved by using either passive or active microrings. If the incoming wavelength is centered on the resonance pic of the filter, the light is coupled into the MR, extracted from input port and transfer to drop port, otherwise it is transferred to through port. This function can be used to construct optical router and switch [21] [45]. An ultra-compact flat-top MR filter based on submicron silicon photonic wire has been demonstrated in [46].

2.1.4 Photodetector

At receiver side, the optical signal is converted back into electrical domain by the photodetector.

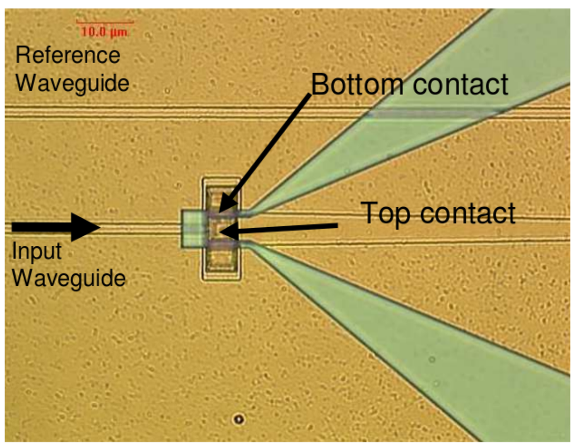


Figure 2.7: Optical microscopy top view of the photodetector integrated at the end of SOI waveguide [6].

Figure 2.7 shows a photodetector integrated at the end of the SOI waveguide. An active material compatible with silicon technology is required for monolithic integration of a photodetector in SOI waveguide [5]. Germanium is hence the preferred option for high performance photodetector due to its shortest absorption length. Vivien *et al.* demonstrated a compact Ge-on-Si pin photodetectors integrated in SOI rib waveguide with improved performances: low dark current (60mA/cm^2), high cut-off frequency (42GHz) and high responsivity (1A/W) at the wavelength of 1.55μm [6]. The photodetector sensitivity (in dBm) is one of most important parameter for energy-efficient communication, which determines the lowest power at which optical signal can be detected.

2.2 Shared Channel Interconnect

2.2.1 Shared Channel Design

A simple photonic interconnect called Single Writer Single Reader (SWSR) design is composed of one writer (transmitter) and one reader (receiver) placed next to a waveguide, called Single Writer Single Reader (SWSR) design. Each pair of source and destination is assigned to a specified channel. This point-to-point connection presents an optimal solution for latency and energy consumption due to the fact that the communication between IP cores are not influenced by each other, but it lacks of scalability, same as a point-to-point electrical interconnection. In this context, shared channel design is proposed to fully exploit photonic interconnect, where multiple readers or/and writers are attached to the same waveguide as shown in figure 2.8, such as Single Writer Multiple Readers (SWMR), Multiple Writers Single Reader (MWSR), and Multiple Writers Multiple Readers (MWMR).

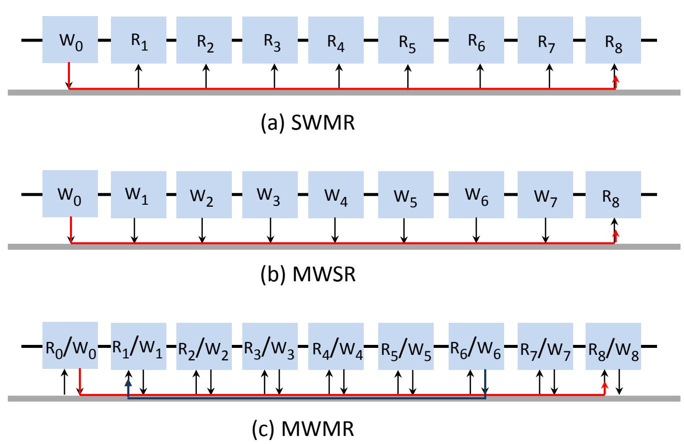


Figure 2.8: Different optical shared channel designs (a) SWMR (b) MWSR (c) MWMR. The black arrow represents the permission of access the channel. The red lines with arrow show the current communication on the channel.

In SWMR design, the writer has a dedicated sending channel to avoid global arbitration and all the readers have access to this channel. However, it is not an energy-efficient solution, since all the data from a core needs to be broadcast to all the others as proposed in ATAC [47]. Another more energy-efficient solution is to employ a local arbitration at writer side before sending data, referred as MWSR channel design. In this case, multiple writers and a single reader are attached on a given channel as proposed in Corona [9]. In both SWMR and MWSR ONoC, the number of nanophotonic components such as MRs, optical interfaces, and waveguides should be proportional to the number of cores, limiting the scalability of the network. MWMR channel is a more scalable design, where each channel is flexible for all writers and readers. The contention at writer and reader side can be resolved by arbitration strategy. MWMR designs achieve better channel utilization and scalability. However, with fully shared resources, many readers and writers are attached to the same waveguide. Then the optical signal may pass through these optical devices suffering from higher optical losses, which increase the optical power emitted by lasers for a target BER.

2.2.2 Wavelength Division Multiplexing

As the channel design with shared single-wavelength waveguides by many readers and writers lacks of scalability, the use of Wavelength Division Multiplexing (WDM), in which multiple wavelengths are transmitted through the same waveguide, can overcome this problem. Indeed, WDM allows multiple signals to be transmitted simultaneously, facilitating higher throughput effectively. Several wavelength-specific lasers are used to ensure support the multiple wavelengths emissions. Moreover, several wavelength-specific MRs are placed on a common waveguide, then a specific wavelength can be extracted from WDM signal at the receiver side. In order to avoid filtering other wavelength channel, MR should have a small ring diameter to create a large FSR, offering important wavelength spacing [48]. Another advantage provided by WDM is the reduction of the excessive contention faced in single-wavelength network. When WDM is used, a physical channel can be divided into many logical independent channels, corresponding to different wavelengths. Thus, the contention possibility will be significantly decreased. However, increasing channel bandwidth requires the use of more MRs and lasers leading to area overhead and more losses.

2.3 WDM-based ONoC Architecture

Current existing WDM-based ONoC architectures can be classified into two categories: all-optical interconnection and combined electrical and optical interconnection.

(1) All-optical interconnection: small-scale ONoC usually uses bus, ring, and crossbar topologies to realise all-optical interconnection. Several passive networks, such as λ-router [7], RPNoC [49] and ORNoC [50] use SWSR optical crossbar and do not require any dynamical reconfigurable switching element. They are WDM-based all-optical point-to-point connecting network. As shown in figure 2.9, a N×N λ-router needs N waveguides, N wavelengths, and multiple basic 2×2 switching elements to achieve fully non-blocking switching. Each node has a dedicated optical channel to every other node. Contention-free communication leads to low latency and high throughput, but it is difficult to implement large-scale architecture due to constraint on the number of wavelengths. Preston *et al.* showed that the maximum number of wavelengths that can be transmitted concurrently in a same waveguide is around 62 in order to satisfy a low crosstalk when assuming 10Gb/s data rate [51].

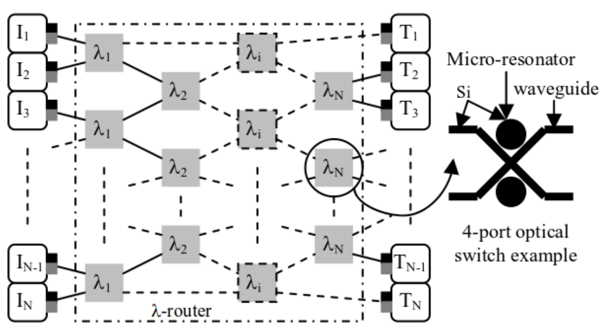


Figure 2.9: (a) N × N λ-router architecture, (b) 4-port optical switch architecture example [7].

Bartolini *et al.* proposed Olympic all-optical architecture using a hierarchical topology made up of ring. Intra-cluster communications use their local optical ring, while inter-cluster communications use the global optical ring [52]. Grani *et al.* proposed an all-optical reconfigurable ONoC architecture to reduce path-setup latency as shown in figure 2.10. A central arbiter is employed to handle set-up request and makes set-up latency not dependent on the number of hops to be traversed. A photonic ring connects each core to the arbiter and an optical folded-torus NoC serves data exchange [8].

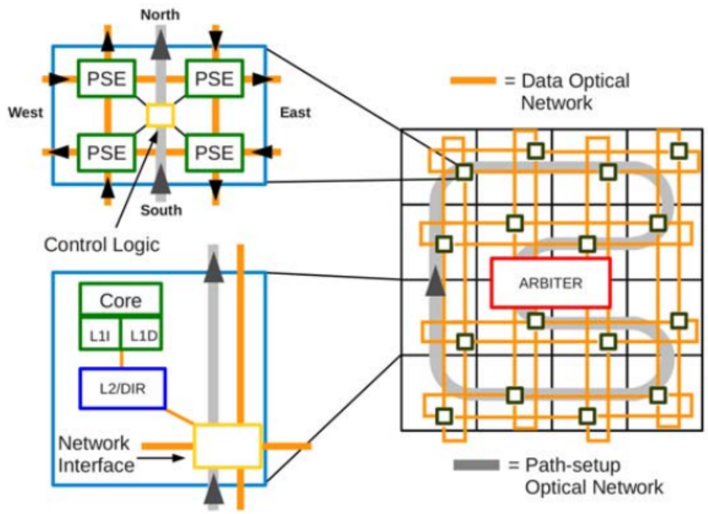


Figure 2.10: A 3D ONoC consists of an photonic ring (in grey) used by the central arbiter for path set-up procedure and an optical folded Torus (in orange) for inter-core communications [8].

(2) Combined electrical and optical interconnection: for large-scale ONoC, optical-electrical hybrid interconnection is employed due to its high network scalability and bandwidth utilization. The optical layer is used for data transfer and the electrical layer for buffering and processing. Vantrease *et al.* proposed Corona architecture implementing MWSR optical crossbar to fully interconnect 64 clusters and electrical link to interconnect 4 cores in a cluster. Corona takes use of optical token-ring arbitration to avoid writing conflict among writers (as shown in figure 2.11) [9]. Each node is equipped with an array of wavelength-specific detectors and can divert one-bit token to obtain grant for the corresponding channel. Meanwhile, the same number of modulators is employed to inject token onto the arbitration waveguide when a transmission is finished.

In order to avoid arbitration at writer side, Kurian *et al.* proposed a SWMR design, called ATAC. In ATAC, each cluster can modulate data on an unique wavelength and receive signals on all other wavelengths. However, broadcasting message to all the readers leads to higher power consumption [47]. Pan *et al.* proposed Firefly architecture that employed Reservation-assisted SWMR (R-SWMR) optical crossbar to avoid power hungry broadcasting. The large crossbar is divided into multiple smaller crossbars to avoid global arbitration and reduce the hardware complexity [53]. Furthermore, they extended Firefly and proposed Flexishare architecture implementing a MWMR optical crossbar. Token stream arbitration is used to resolve contention at writer side. The channel utilization is improved compared to token ring arbitration due to the injection of a set of tokens. At reader side, the same mechanism as reservation-assisted scheme in Firfly is adopted to activate the corresponding destination core [54]. Xu *et al.* proposed "channel borrowing" optical crossbar architecture. The control and latency overhead of arbitration are much lower due to the greatly reduced sharing degree. Besides, the dependency between the number of microring resonators and the radix of crossbar is removed in the proposed crossbar [55].

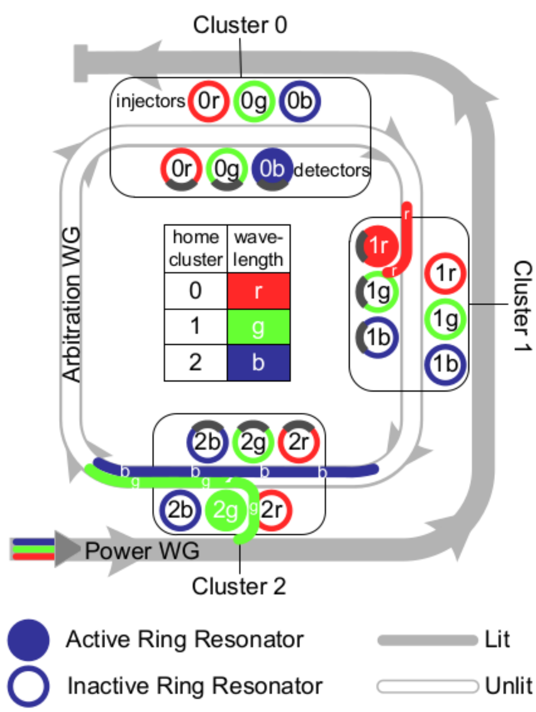


Figure 2.11: Token ring arbitration in Corona [9].

Shacham *et al.* proposed a hybrid ONoC that uses an optical circuit-switched network to transmit high bandwidth messages, and an electrical packet-switched network to transmit small messages and control packets [56]. A path setup process is required before each data transmission. Once the photonic path is established, the data are transmitted to the destination node without buffering. To further improve channel efficiency, Le Beux *et al.* proposed an architecture named Chameleon [57]. The reconfigurability feature allows to open and close dedicated channels between IP cores.

Hence, the bandwidth is shared according to the communication requirements, thus leading to a highly reusable network. Besides, the ring topology and the regular layout of interfaces contribute to a good scalability of Chameleon. The wavelength allocation in Chameleon can be achieved at run-time, through a communication protocol [58], or specified at compile time [59], using static mapping method. The configurations generated by the later method lead to networks characterized by a low latency (no arbitration is required), highly energy-efficient (dedicated channels are used). Similarly, Wu *et al.* presented a Sectioned Undirectional Optical Ring (SUOR) NoC as illustrated in figure 2.12 [10]. The set-up control subsystem is composed of cluster agents placed in the center of chip in order to shorten set-up overhead. Moreover, the clusters are connected with optical waveguides which are aligned as closed-loops on the chip. In their design, a single waveguide can support multiple concurrent transactions dynamically based on the arbitration, improving the network throughput.

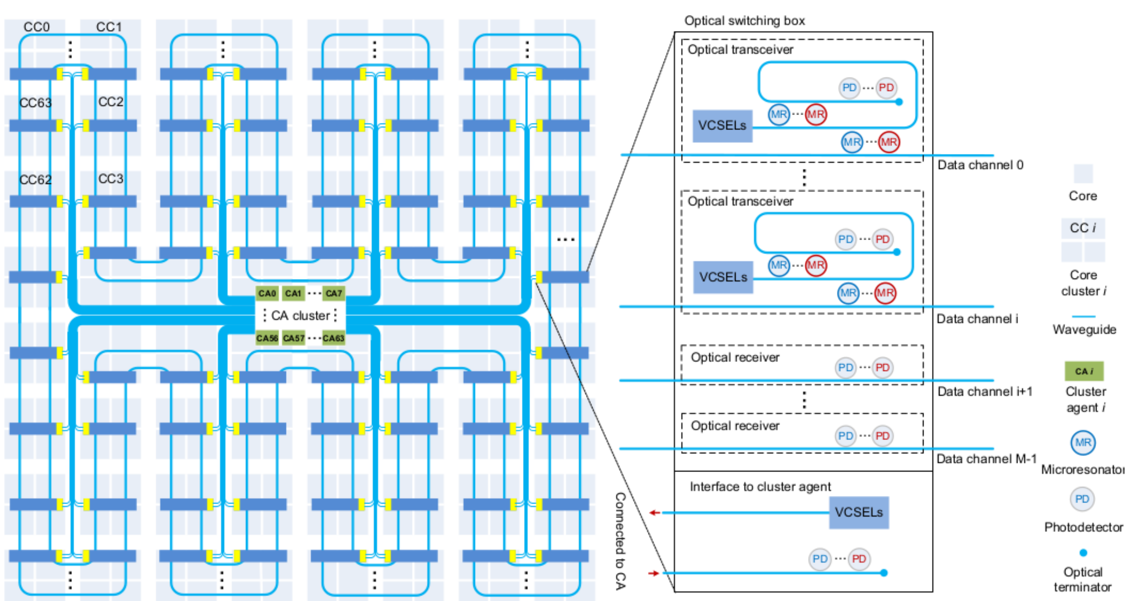


Figure 2.12: The overview of SUOR architecture and its floorplan [10].

Zhang *et al.* proposed a centralized controlled ONoC architecture, referred as Ouroboros Network (ON). Space-Division-Multiple-Access (SDMA) is employed to allow multiple accesses of the optical network simultaneously. However, the pre-knowledge of each subtask should be known by the control center [60].

Chen *et al.* [61] proposed a new hierarchical ONoC structure named CWNoC. The whole electrical control network is divided into several central-controlled subnets, and each subnet employs a centralized control mechanism. Hence the set-up overhead is ultra-low and independent on the network diameter. A low-power fat tree-based optical network-on-chip, named FONoC, is proposed in [62]. FONoC does not require building a separate electronic NoC for network control. It carries both payload data and network control data on the same optical network.

Finally, the insertion loss and crosstalk issues are addressed in recent communication architecture. Fusella *et al.* proposed to employ a hybrid-topology in the photonic layer as a solution to reduce the insertion loss and crosstalk, hence improving the network scalability, increasing the number of available wavelengths, and reducing the energy consumption [63].

2.4 BER Estimation

In this part, we describe how to calculate the power consumption for the application-dependant transmissions between source and destination IPs running on ONoC architecture, taking into account the states of the crossed MRs.

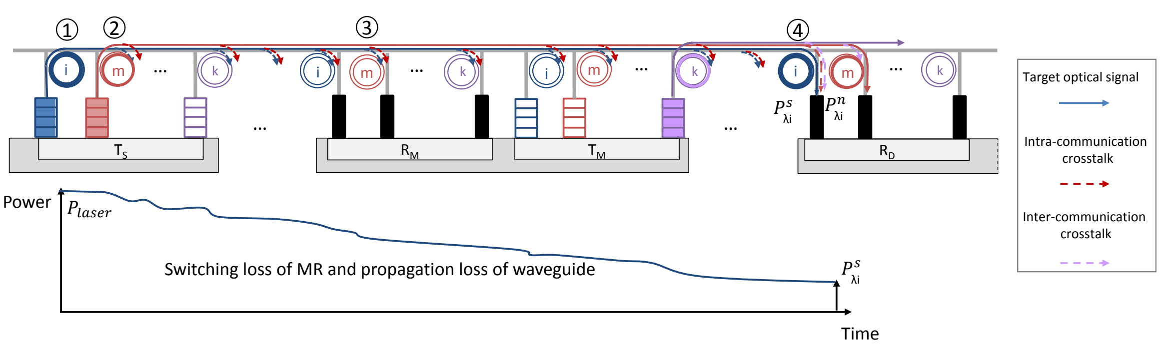


Figure 2.13: Signal propagation in a generic WDM-based channel.

Figure 2.13 illustrates a generic WDM-based channel propagating optical signals. N_λ wavelengths can be used for communicating within the WDM-based shared channel. As shown in figure 2.13, a communication between source transmitter (T_S) and destination receiver (R_D) can be established by using the assigned channels λ_i (blue) and λ_m (red) through on-chip lasers and MRs. Here, we focus on the optical signal on

λ_i. It is injected into the waveguide by the on-chip laser specified to λ_i (mark ①) in transmitter T$_S$. Then the optical signal propagates through the intermediate blocks until reaching destination receiver R$_D$. It suffers from switching losses of MRs and propagation loss of waveguide. In particular, when it passes through other transmitters and receivers. As stated in section 2.1.3, the signal transmission of MR$_{\lambda_m}$ depends on its state: OFF-state and ON-state leading to $\lambda_m + \Delta\lambda$ and λ_m resonant wavelengths respectively (see figure 2.6). Regarding the losses of a wavelength crossing a MR, we consider three resonances in our models: the main at λ_m, and two mores at $\lambda_m + FSR$ and $\lambda_m - FSR$, which corresponds to the two successive resonances. Hence, on the through port, the signal power $P_{T,Mon_{\lambda_m}}[i]$ at λ_i for a MR specific to λ_m in the ON-state (mark ②) is defined as:

$$P_{T,Mon_{\lambda_m}}[i] \quad = \quad P_{input}[i] + \Phi^t_{dB}(\lambda_i, \lambda_m - FSR) + \Phi^t_{dB}(\lambda_i, \lambda_m) + \Phi^t_{dB}(\lambda_i, \lambda_m + FSR) \tag{2.5}$$

$P_{input}[i]$ is the input optical power of signal at λ_i. Φ^t is the transmission function on through port as described in Eq. 2.4. Then the signal power $P_{T,Moff_{\lambda_m}}[i]$ at λ_i for a MR specific to λ_m in the OFF-state (mark ③) is defined as

$$\begin{aligned} P_{T,Moff_{\lambda_m}}[i] \quad = \quad & P_{input}[i] + \Phi^t_{dB}(\lambda_i, \lambda_m + \Delta\lambda - FSR) + \Phi^t_{dB}(\lambda_i, \lambda_m + \Delta\lambda) + ... \\ & \Phi^t_{dB}(\lambda_i, \lambda_m + \Delta\lambda + FSR) \end{aligned} \tag{2.6}$$

Similarly, the transmissions on the drop ports, i.e. $P_{D,Mon_{\lambda_m}}[i]$ and $P_{D,Moff_{\lambda_m}}[i]$ are obtained similarly.

$$P_{D,Mon_{\lambda_m}}[i] \quad = \quad P_{input}[i] + \Phi^d_{dB}(\lambda_i, \lambda_m - FSR) + \Phi^d_{dB}(\lambda_i, \lambda_m) + \Phi^d_{dB}(\lambda_i, \lambda_m + FSR) \tag{2.7}$$

$$P_{D,Moff_{\lambda_m}}[i] \quad = \quad P_{input}[i] + \Phi^d_{dB}(\lambda_i, \lambda_m - FSR) + \Phi^d_{dB}(\lambda_i, \lambda_m) + \Phi^d_{dB}(\lambda_i, \lambda_m + FSR) \tag{2.8}$$

At receiver side, the optical signal on λ_i is ejected from waveguide. In order to calculate the power at receiver side, we estimate the losses experienced by the all the transmitting components, from laser source to destination, taking into account the states of the crossed MRs. This leads to Eq. 2.9 which corresponds to the received signal power ($P^s_{\lambda_i}$) on a photodetector connected to a MR at wavelength λ_i.

$$P^s_{\lambda_i} \quad = \quad P_{laser} + L_{M,on}[i] + L_{M,off}[i] + L_P[i] \tag{2.9}$$

P_{laser} is the optical laser output power. $L_{M,on}[j]$ and $L_{M,off}[j]$ are the losses experienced by signals due to OFF-state and ON-state MRs crossing. $L_P[j]$ is the propagation loss along the path.

Please note that figure 2.13 shows two fundamentally sources of crosstalk: inter-communication and intra-communication crosstalks. The first one comes from the undesirable coupling from other wavelengths allocated for the same transmission (e.g. between λ_i and λ_j). The second one occurs when two different transmissions share the same segment of waveguide simultaneously (e.g.between λ_i and λ_k). Hence, the total crosstalk power at wavelength λ_i can be calculated as:

$$P^n_{\lambda_i} \quad = \quad \sum^{j \in \{m,k\}} P_{laser} + L_{MR,on}[j] + L_{MR,off}[j] + L_P[j] \tag{2.10}$$

Then a simplified SNR model for the input of the photodetector of λ_i can be defined by

$$SNR_{\lambda_i} \quad = \quad \frac{P^s_{\lambda_i}}{P^n_{\lambda_i} + P^n_{pd}} \tag{2.11}$$

Where P^n_{pd} is the noise added by the photodetector. Finally, the BER is obtained by Eq. 2.12.

$$BER = \frac{1}{2} erfc\left(\frac{SNR}{2\sqrt{2}}\right) \tag{2.12}$$

Where erfc is the complementary error function. From the above analysis, we can see that BER deeply depends on the number of co-propagating wavelengths. To ensure efficient communication, the objective is to minimize the BER.

2.5 Energy Efficiency and Communication Performance Challenges in ONoC

Silicon photonic interconnect can offer advantages such as low transmission latency and high bandwidth. However, photonics devices introduce new electro-magnetic effects (insertion loss and crosstalk), having no electronic equivalent, that greatly impact the energy efficiency of ONoC [31]. Besides, lasers are the major source of ONoC power consumption. Therefore, in this thesis, we focus on the improvement of the energy efficiency of photonic interconnect by adapting laser power, taking into account crosstalk and target communication BER. Furthermore, the trade-off between energy efficiency and communication performance is also investigated.

2.5.1 Crosstalk Noise

Crosstalk noise is introduced by the interference of different signals degrading the reliability of networks. The standard design-time device-level solutions are proposed to reduce crosstalk in photonic interconnect. For example, Sanchis *et al.* proposed to minimize crosstalk in SOI waveguides by choosing the optimum crossing angle. Experimental results showed that the optimum crossing angle is not 90 degree but either 60 or 120 degrees [64]. Related to microring resonator, Jayatilleka *et al.* investigated crosstalk of first- and second-order MR. They demonstrates that the crosstalk noise can be reduced by increasing the order of MR [65].

The crosstalk noise at device-level seems negligible, but the accumulated crosstalk at network-level may severely affect SNR performance of network. Some researches dedicated to explore the crosstalk noise on network-level have been done. Xie *et al.* studied the crosstalk noise, SNR and BER of mesh-based network. Crux optical router is proposed to minimize insertion loss and crosstalk [66]. Nikdast *et al.* analysed the crosstalk noise and SNR in folded-torus-based [67] and fat-tree-based ONoC [68]. These studies in literature above are generally focusing the worst-case scenario in terms of demands on technology. Fusella *et al.* proposed a mapping tool to mitigate the crosstalk by reducing the number of communications sharing a waveguide [69]. These aforementioned models use only a single-wavelength approach and the crosstalk originates from only one wavelength. However, the SNR is reversely related to the number of wavelengths employed in the network and the crosstalk issue is more critical in WDM ONoC. Duong *et al.* studied crosstalk noise in a Dense Wavelength Division Multiplexing (DWDM) ring-based ONoC. They demonstrated that a higher Q-factor can reduce crosstalk, in turn improve energy efficiency [70]. Chan *et al.* proposed a optical network simulator, called PhoenixSim [48], to further evaluate the photonic network design space. They presented a methodology to jointly explore the physical-level and system-level designs of the network topologies [38]. These proposals study power consumption for several existing topologies, but do not use that information to reduce energy consumption of network.

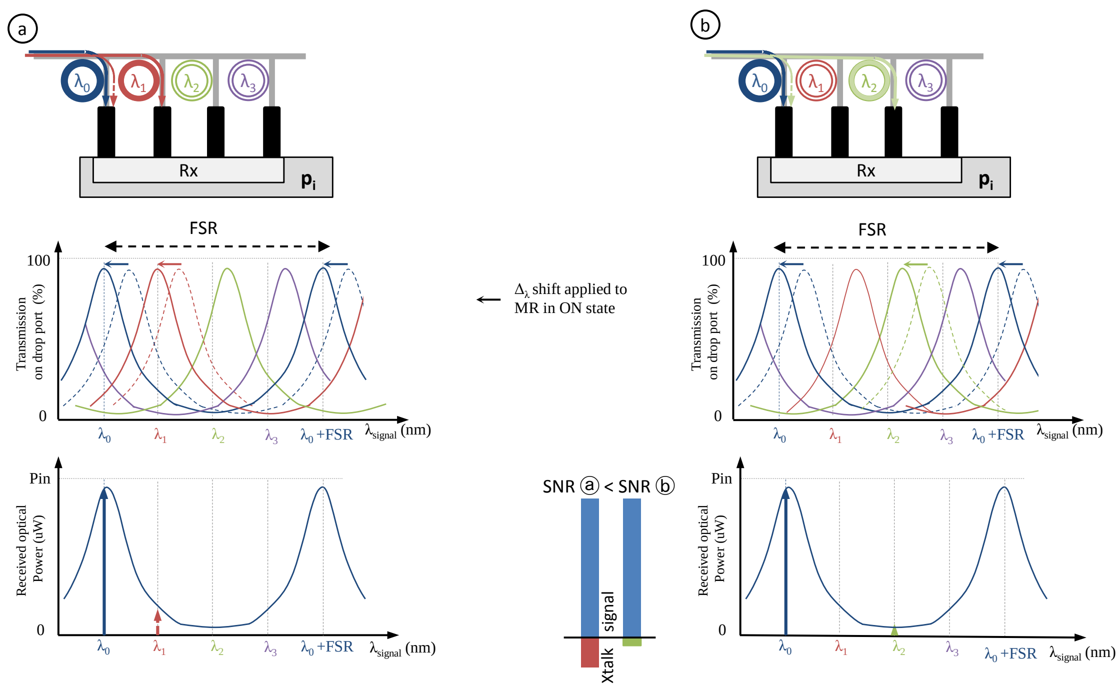


Figure 2.14: Receiver: the loss due to crosstalk depends on the wavelengths used to communicate, (a) two neighboring wavelengths (b) two distinct wavelengths.

The crosstalk power received by the photodetector depends on the MR drop spectrum: the closer the other signals to the resonant wavelength, the higher the crosstalk. Figure 2.14 illustrates an interface receiver for a 4-wavelength ONoC example. For case ⓐ, we assume that the interface is receiving two signals on wavelengths λ_0 (blue) and λ_1 (red), i.e. blue and red MRs are set to the ON state to ensure the corresponding wavelengths from waveguide. Green and Purple MR are OFF. For this purpose, voltage tuning is applied to shift by $\Delta\lambda$ nm the resonant wavelength of MRs and then aligned with the signal wavelengths. Blue and red signals are thus dropped from the horizontal waveguide and reach photodetectors. However, blue MR also drops a part of the red signal (dashed red arrow in the figure), which leads to crosstalk noise. For case ⓐ, the crosstalk is potentially high since the signals are transmitted using adjacent wavelengths [71]. To reduce the crosstalk, one can communicate using higher spacing between the wavelengths. For case ⓑ, we select wavelength λ_0 and λ_2 (green), which leads to a lower crosstalk on the photodetector and thus it can offer the possibility to reduce the power of the emitted laser. Chittamuru *et al.* propose a crosstalk mitigation technique to increase channel spacing between adjacent wavelengths in DWDM [72]. However, this wavelength spacing technique leads to the reduction of DWDM degree in the waveguide, in turn reduces channel bandwidth and network throughput.

2.5.2 Laser Power Management

Managing an ONoC to execute a given application is a tedious task, especially when power consumption is closely related to BER objectives. Lasers power consumption is an important part of ONoC static power consumption. To improve energy efficiency, some approaches presented in the literature propose to tune the laser power. Zhou *et al.* proposed a prediction strategy based on network usage to modulate the laser power: the laser power is adjusted according to predictions based on previous activity [73]. They designed an efficient scheme with a bandwidth traffic predictor to increase the channel utilization and reduce static optical power in the external laser. On-chip lasers are more static power saving compared to off-chip laser source. Since on-chip laser can be powered off when there is no data transfer, this would significantly reduce power consumption if the application load is not high. In [74], a feedback loop configures the laser power in a linear mode by using a 6-bit DAC. However, this solution suffers from a poor scalability due to the need to estimate the BER on the receiver side and to the feedback control. Furthermore, there is no analysis of the penalty induced by the BER estimation. In general, laser power dissipation depends on the bandwidth of data channel and photonic losses along optical path. In order to accurately reduce power dissipation, Chen *et al.* proposed an on-off control strategy to maximize the energy efficiency [75]. The optical links can access memory (L2 caches) and can be activated/deactivated at run-time according to the size of application's working set. According to simulation results, on-chip laser's average power consumption is reduced by 23.8% using on-off control strategy compared to the strategy that always keeps lasers on. This method can be efficiently combined to the laser sharing and placing strategy they propose to improve the ONoC energy efficiency [76]. In [77], a CMOS driver is designed to configure the laser power to OFF, standby, mid- and full-power modes. The laser power is switched among the four modes to provide just-enough optical power for the desired communication bandwidth and for target BER. The higher the communication bandwidth, the higher the laser power. They also consider a time and power penalty for OFF-mode and standby-mode respectively. Ortin-Obon *et al.* proposed a tool to optimize laser power for wavelength-routed optical ring communication matrix that a specific waveguide-wavelength pair is allocated for each communication [78]. However, this tool is only used for wavelength-routed ONoC which is limited to small-scale architectures.

2.5.3 Communication Performance

As we know, the use of WDM can open high bandwidth channel, leading to a small transmission time. Figure 2.15 illustrates several configurations for the some communication between p_1 and p_2. The first case, figure 2.15(a) illustrates a communication requiring 8 clock cycles if the communication is performed on 1 wavelength. The higher the number of reserved wavelength for one communication, the smaller the transmission time. Hence, when 2 wavelengths are reserved, the transmission time is reduced to 4 clock cycles as shown in figure 2.15(b). Specially, the smallest transmission time is reached (2 clock cycles) when all free wavelengths are used for this communication as shown in figure 2.15(d). However, in this best case, there are no more free wavelengths along this optical path for other communications, i.e. communication between p_0 and p_3 can not be established as shown in figure 2.15(e). Besides, as previously mentioned, the laser power consumption is proportional to the communication bandwidth. This leads to the following conflicting objectives: high performance communication tend to rely on an exhaustive use of the available wavelengths while energy efficient communication involve a parsimonious use of wavelengths.

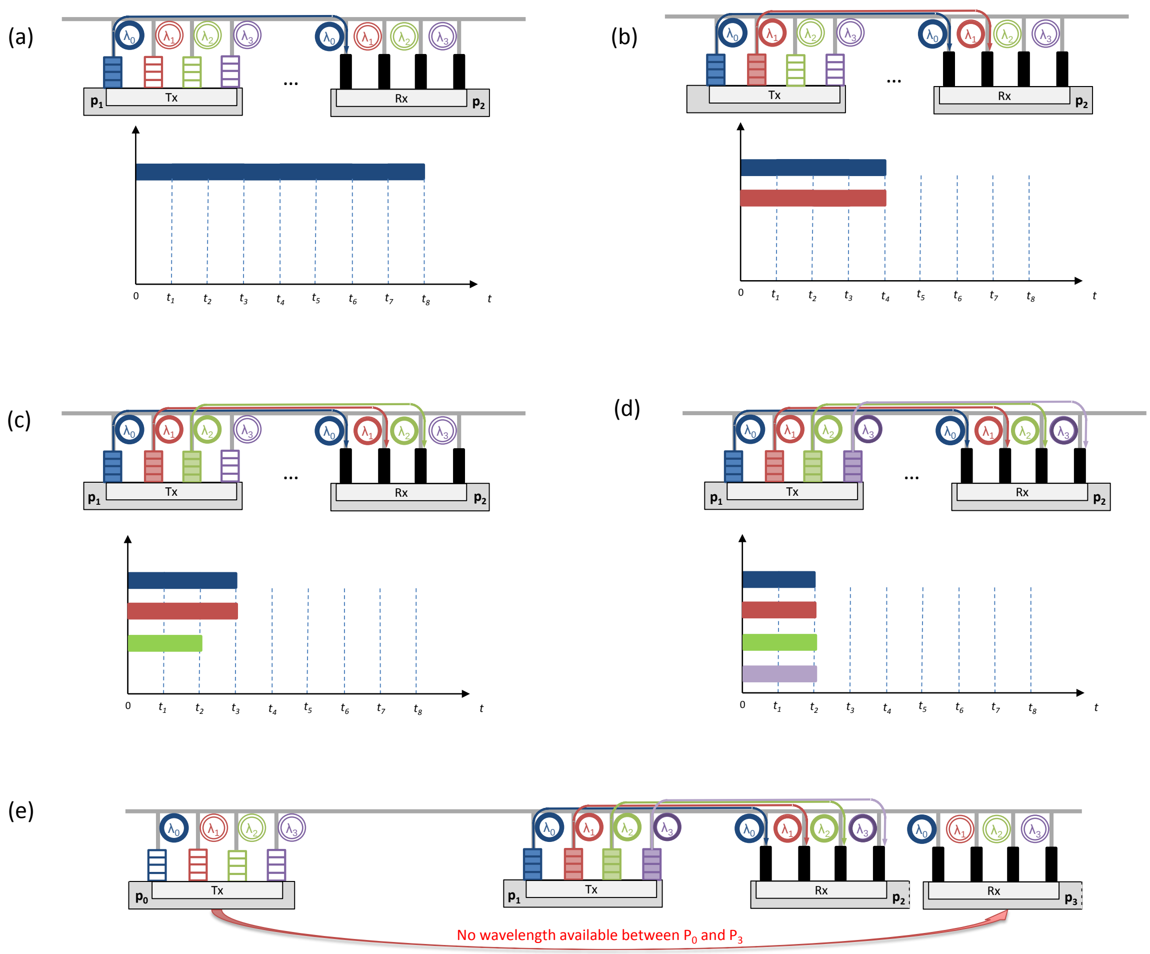


Figure 2.15: Wavelength allocation impact on communication performance: (a) 1 wavelength (b) 2 wavelengths (c) 3 wavelengths (d) 4 wavelengths for communication between p_1 and p_2. (e) No wavelength available between p_0 and p_3.

2.6 Discussions

The work published in the literature address several interesting aspects of ONoC, but none of them propose a global approach to ensure wavelength allocation, laser management, and offline-online management. Considering this state of the art, this thesis proposes several contributions.

The main contribution of this thesis is to allocate one or a set of wavelengths for each communication, jointly consider the most appropriate position of each reserved wavelength to improve network energy efficiency. The wavelength allocation issue is addressed by offline and online approach respectively.

In offline approach, we combine wavelength allocation and laser power scaling to find the trade-off of the performance and energy consumption. The methodology is formed as a framework handling several inputs, such as application, architecture, and device parameters. On the one hand, the exploration of design space at both device level and system level is allowed; on the other hand, a set of performance and energy efficiency trade-offs can be generated once the input parameters are defined. The resulting Pareto solutions include 1) low-power solutions, which tend to minimize the number of used wavelengths, and 2) high-performance solutions, for which multiple wavelengths are allocated to shorten the communication time. Finally, these solutions can be embedded in a controller and deployed at run-time according to execution requirement (e.g. high performance and low power).

However, the offline approach is restricted to a predefined application, i.e. it is not able to handle a dynamic workload and run-time deployment of tasks. In this case, dynamic policies to allocate the optical resources at run-time in a MPSoC architectures should be considered for that objective. We propose a tree-based electrical set-up control network, named TbNoC for online management of ONoC. We partition the global electrical control network into several hierarchical central-controlled agents and centralize the resource arbitration for purpose of reducing set-up path latency. Then, a minimization of crosstalk algorithm is employed to separate the wavelengths for optical signals to reduce crosstalk, hence, the energy efficiency is improved. Finally, the fully shared resource mechanism is more flexible compared to the limited access rules implemented in Corona and SUOR, but it results in a higher complexity of network. Consequently, a trade-off between flexibility and complexity of network should be considered by the designer.

In our thesis, the wavelength allocation can be specified at compile time or achieved at run-time. In the former case, the communication requests are known in advance, then an extra offline simulation is required to select proper wavelengths. While in the latter case, a sequence of communication requests is received at run-time and an optical path is set-up in response to each request. Each approach is suitable for different application scenarios and both methods could be jointly used in ONoC.

2.7 Conclusion

We have introduced related preliminary knowledge of on-chip silicon photonic interconnects in this chapter, including photonic interconnect basics and WDM-based ONoC architecture etc. In WDM interconnect, wavelength allocation plays an important role since it allows linking application level communications to resource allocation at the hardware level. The multiples wavelengths in a shared waveguide can offer high bandwidth channel. However, optical signals of various wavelengths can interfere with each other through optical switching elements, introducing crosstalk noise. Thus wavelength allocation as a technique to improve communication performance and energy efficiency can be addressed. In the following chapters, the proposed offline and online wavelength allocation approaches will be detailed.

Chapter 3

Offline Optimization of Performance and Energy for Online Bandwidth and Laser Power Configuration

ONoCs are based on integrated waveguides, carrying optical signals, and optical devices allowing to inject or drop the signals into these waveguides from an electrical interface. This technology offers low latency and high bandwidth properties. The waveguide can be shared by multiple senders and receivers when DWDM is used to support multiple transactions simultaneously, providing a significant increase in bandwidth. Indeed, several communications between different couples of source and destination cores can be established in parallel by using different wavelengths. However, simultaneous transmissions, on close adjacent wavelengths, may introduce inter-channel crosstalk noise through different optical switching elements within the network [70], which negatively impacts the SNR and, therefore, the BER. The BER to be reached depends on the application requirements and is closely related to the optical devices characteristics: it depends on the photodetector sensitivity [31], on the MR transmission spectrum, and on the optical signal power emitted by the laser sources. The higher the number of signals propagating simultaneously, the higher the crosstalk and the higher the laser output power needed. This leads to the following conflicting objectives: high performance communications tend to rely on an exhaustive use of the available wavelengths while energy efficient communications involve a parsimonious use of signals which occupy distinct and separate wavelengths. Using an ONoC to execute a given application is thus a tedious task, especially if performance, power, and BER objectives are likely to evolve with the execution context.

This chapter addresses this problem by proposing an offline methodology allowing the exploration of performances and energy consumption trade-off for ONoC-based MPSoC. In order to find the best trade-off, we combine Wavelength Allocation (WA) and laser power scaling. The key ideas are to i) allocate just enough wavelengths for each communication to reduce the application execution time and ii) tune the laser driver current to provide just-enough optical power to meet the application BER requirements. The multi-objective optimization problem is solved by a proposed framework handling several inputs, such as application, architecture, and device parameters etc. On the one hand, the exploration of design space at both device level and system level is allowed; on the other hand, a set of performance and energy efficiency trade-offs can be generated once the input parameters are defined. From the set of Pareto points, the designer can select the most appropriate solutions at run-time according to the specific requirement. In this chapter, we first present key concept of wavelength allocation and laser power configuration in Section 3.1. Then, the proposed methodology is illustrated in Section 3.2. The problem formalization and resolution are presented in Section 3.3 and Section 3.4 respectively. Section 3.5 introduces the optical network interface management. Finally, the conclusion of this chapter is given in Section 3.6.

3.1 Key Concept of Wavelength Allocation and Laser Power Configuration

The aim of this work is to identify, at design time, a set of ONoC configuration modes that will be embedded in a system and deployed at run-time. Figure 3.1 illustrates the key concepts we investigate. Applications are represented as a Directed Acyclic Graph (DAG), where each vertex t_i represents a task and each directed edge ($c_{i\to j}$) represents a communication from t_i to t_j. In this example, we assume that tasks t_0, t_1, and t_2 are mapped onto processors p_0, p_1, and p_2, respectively. Communications $c_{0\to 1}$, $c_{0\to 2}$, and $c_{1\to 2}$ are implemented using an ONoC which is configured according to execution performance and energy requirements. WDM-based optical interconnect allows to configure the wavelengths allocated to communications and the laser output power. The following illustrates four configuration modes:

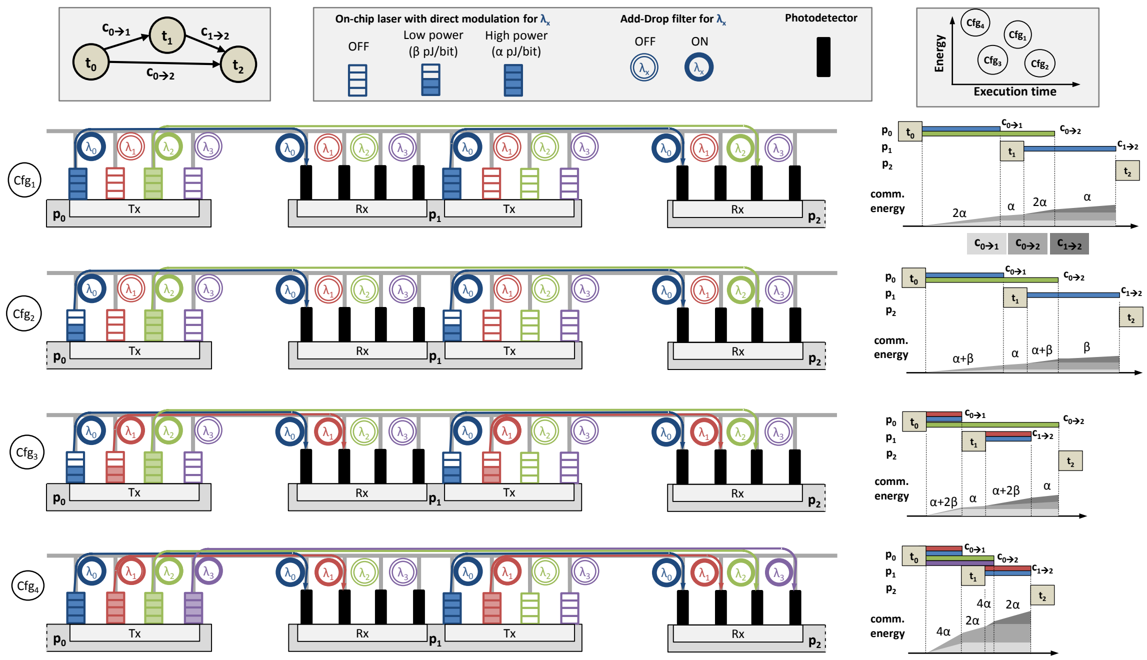


Figure 3.1: WDM-based optical interconnect allows for configuring the wavelengths allocated to communications and the laser output power.

- **Baseline mode** Cfg_1. In this mode, a single wavelength is allocated to each communication and the activated lasers are configured to emit high power optical signals. In the figure, lasers at wavelengths λ_0 (blue) and λ_2 (green) are allocated to $c_{0\rightarrow1}$ and $c_{0\rightarrow2}$ respectively. We assume a direct ON-OFF Keying (OOK) modulation, i.e. no modulator is needed. The signals are injected into a waveguide using MRs tuned to align their resonant wavelengths with signal wavelengths (ON state). The signals propagate along the waveguide until they reach their destination: signals at wavelength λ_0 and λ_2 are ejected in p_1 and p_2 interfaces respectively (the corresponding MRs in the interfaces are set to the ON state). Since λ_0 is used only from p_0 to p_1, it is reused to implement $c_{1\rightarrow2}$ at the next communication time, thus maximizing the wavelength occupation in the network. In this example, we assume that the targeted BER is reached for the all communications. The right hand-side of the figure illustrates the task and communication schedule: p_1 starts executing t_1 once all the data have been received from p_0 and then send processed data to p_2. p_2 first receives data from p_0 but the execution of t_2 is delayed until data from p_1 have been received. We also represent the ONoC communication energy under the task schedule: since the lasers are configured to output the maximum optical power, each communication consumes the maximum energy per bit, defined as α pJ/bit.

- **Low power mode** Cfg_2 leads to the same communication scheme previously described. However, we assume that a lower optical power emitted by the lasers is enough to reach the targeted BER for $c_{0\to1}$ and $c_{1\to2}$ (1 hop each), in this case, the laser power is fixed at β pJ/bit, while $c_{0\to2}$ still needs a high optical power due to the high losses experienced by the signal (2 hops). While the lower optical power for $c_{0\to1}$ and $c_{1\to2}$ does not impact the task schedule, the energy consumption is reduced since it depends on both α and β pJ/bit (obviously, $\alpha > \beta$).

- **Intermediate mode** Cfg_3 implies a higher bandwidth allocated to $c_{0\to1}$ and $c_{1\to2}$. Indeed, both communications are implemented using λ_0 and λ_1 (red) which leads to 50% reduction in the communication latency. Despite the three simultaneous wavelengths on the waveguide between p_0 and p_2 and hence the possibly higher crosstalk, we assume that the BER requirement is still met when λ_0 and λ_1 lasers are configured to their low-power mode. As depicted on the figure, t_1 and t_2 start executing earlier, which allows for reducing the execution time compared to modes Cfg_1 and Cfg_2. However, one can expect a higher increase of communication energy due to the laser thresholds current and voltage (i.e., $2 \times \beta > \alpha$).

- **High performance mode** Cfg_4 leads to a higher bandwidth for $c_{0\to2}$ since both λ_2 and λ_3 (purple) are allocated to its implementation. Due to the high propagation distance (2 hops), λ_3 laser is also set to the high power mode (α pJ/bit). Furthermore, this mode implies four simultaneous communications, which severely increases the crosstalk. This calls for a high optical power on all the communications, i.e. all the activated lasers are set to the high power mode, even those experiencing relatively low losses (1 hop). Hence, as depicted on the task schedule, the higher bandwidth for $c_{0\to2}$ allows for reducing the application execution time but significantly increases the communication energy.

As depicted on the top-right of the figure, configuration modes Cfg_2, Cfg_3, and Cfg_4 belong to a Pareto front; the suitability of each configuration mode thus depends on the execution context (e.g. power and real time constraints). The network configuration modes are handled by a configuration manager that adapts ONoC configuration mode to satisfy application and execution context requirements on run-time. For instance, a constraint may be an execution deadline for real time systems, the manager thus loads high performance mode. Then the battery of system is nearly ran out, the manager deploys low power mode for purpose of energy saving.

3.2 Proposed Methodology

In this section, we present the proposed framework, which aims at optimizing the ONoC configuration to execute a given application. Figure 3.2 illustrates the framework generating ONoC configuration modes according to user specifications. The flow takes as inputs an application mapped onto a 3D architecture. The application is modeled as a task graph characterized by task execution times, amount of data transmitted between tasks, and minimum BER to be reached. The architecture includes an ONoC implemented on top of processing cores and characterized by a topology, a number of wavelengths and waveguides. The ONoC allows for cores to communicate with each other using optical signals, which is achieved using E/O and O/E conversion. The interfaces are crossed by waveguides propagating the optical signals using WDM in both clockwise and counter-clockwise directions which reduces the maximum communication distance. The mapping of the tasks on the cores gives the communications in the ONoC. The design flow relies on device parameters since they impact the performance of optical communications. Instance of parameters are photodetectors sensitivity, waveguide losses and MR model. Regarding the laser, we take the data-rate, the efficiency, the maximum output power, and the number of power levels available into account.

The aim of the flow is to optimize both power consumption and application execution time. For this purpose, we explore both device-level and system-level parameters. Based on a set of device and system input parameters, a multi-objective optimization is carried out using a genetic algorithm due to the two contradictory objectives. In our genetic algorithm, the ONoC configuration modes are represented by chromosomes and the genes encode both wavelength allocations and laser power levels. Then a simulation is

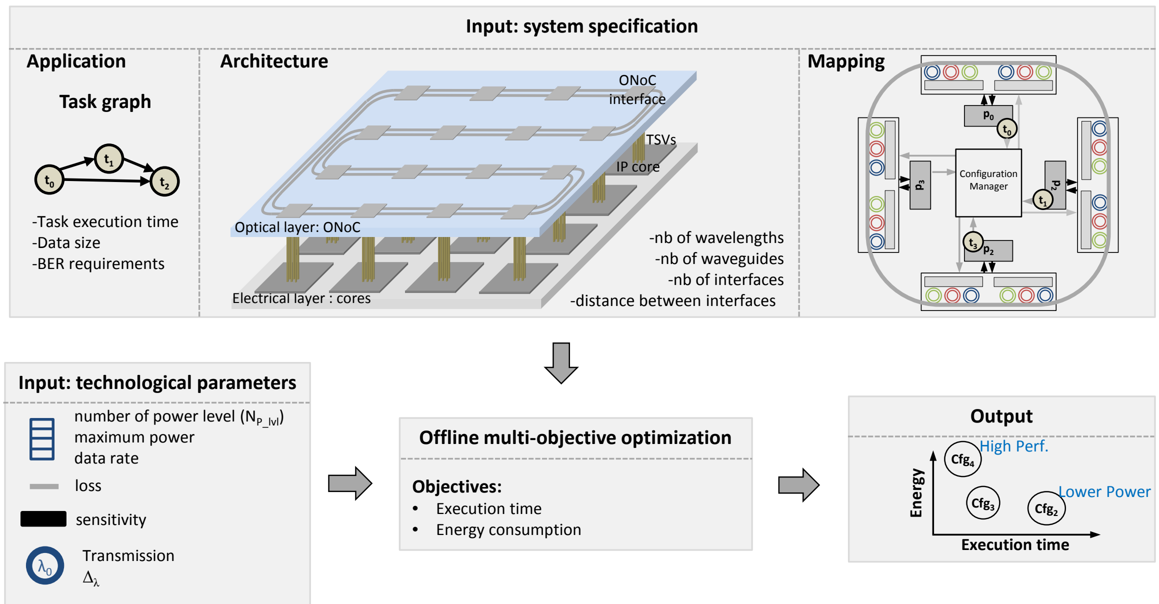


Figure 3.2: Multi-objective design space exploration.

carried out to estimate application execution time and total ONoC energy consumption (i.e. the algorithm fitness functions). Finally, the resulting ONoC configuration modes on a Pareto front from the multi-objectives optimization are thus reported, including low-power solutions, which tends to minimize the number of used wavelengths, and high-performance solutions, for which multiple wavelengths are allocated to shorten the communication time. Then the obtained trade-off configuration modes could be embedded in the system and be loaded on run-time according to the execution context (e.g. high performance and low power) which is out of the scope of this thesis. In summarize, our proposed framework allows not only the exploration of design space in terms of various input parameters, but also the generation of performance and energy efficiency trade-off based on a set of defined input parameters.

3.3 Problem Formalization

In this section, we detail the architecture, application, time, and energy models.

3.3.1 Architecture

Our methodology is implemented on the ring-based ONoC characterized by the number of wavelengths and waveguides. There are not restrictions on the architecture, and the elements we communicate can be simple cores, large clusters or other components. Figure 3.3 shows how a ring-based ONoC architecture is defined. It is composed of two layers interconnected by TSV: i) an electrical layer implementing $n \times n$ IP cores (4×4 in figure 3.3) and ii) an optical layer integrating the ONoC. To connect IP cores to the waveguide, an Optical Network Interface (ONI) is implemented for each IP core. Each ONI integrates a receiver and a transmitter.

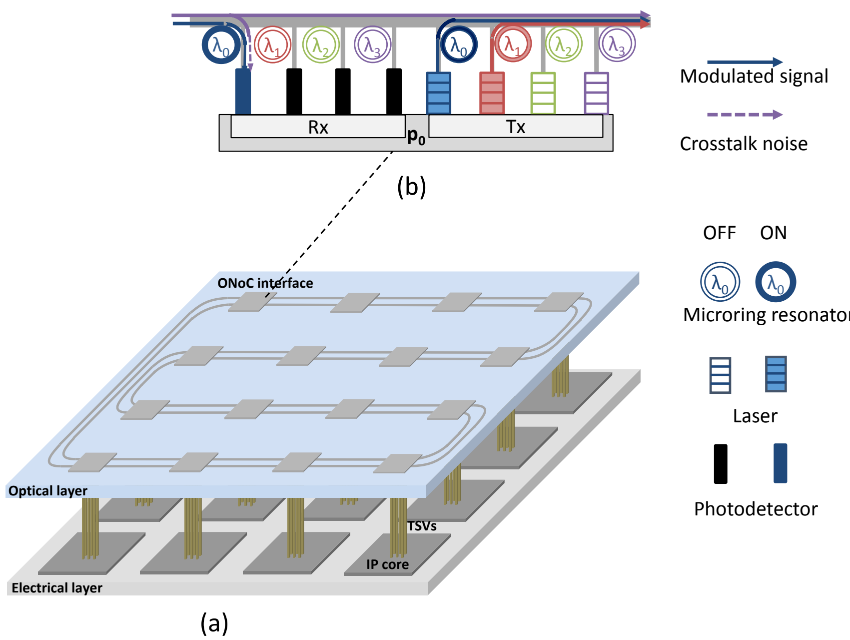


Figure 3.3: Architecture overview: a) ring-based ONoC architecture b) optical network interface.

The transmitter is composed of on-chip laser sources that each on them can emit and inject optical signal at a specific wavelength into the waveguide. The data are directly transmitted from these lasers through current modulation (OOK) and each laser source can also be turned OFF for energy saving. The receiver part includes wavelength-specific MR that can be turned ON or OFF to configure respectively drop (receive) or pass-through operations of the signals at a corresponding wavelength. Signal dropped from a waveguide reaches a photodetector, where opto-electric data conversion generates an electrical signal suitable for the electronics part of the receiver. The considered architecture allows the reuse of wavelengths to realize multiple independent communications in a single waveguide.

Each ONI is crossed by two waveguides propagating optical signals. The waveguide implements a DWDM technique with N_λ wavelengths (λ_0, $\lambda_1 ... \lambda_{N_\lambda - 1}$). Each waveguide is single directional and two waveguides can be used to propagate optical signals in both clockwise and counter-clockwise direction. For each pair of source (p_i) and destination (p_j), we need to decide which direction to take. In order to minimize the power loss along the paths, only the shortest distance is taken into account for each communication between two nodes as described in Eq. 5.2.

$$p_i \rightarrow p_j = \begin{cases} Clockwise & if\ ((i < j) \wedge (|i-j| \leq \frac{N_p}{2})) \vee ((i > j) \wedge (|i-j| \geq \frac{N_p}{2})) \\ Counter-clockwise & otherwise \end{cases} \tag{3.1}$$

With N_p the number of IP cores in the architecture.

3.3.2 Application and Time Model

Definition 1: A Task Graph TG $= G(T, C)$ is a directed graph, where each vertex $t_i \in \mathcal{T}$ represents one task. Each directed edge $c_{i \rightarrow j} = (t_i, t_j) \in C$ characterizes the communication from task t_i to task t_j. The weight of edge is the communication volume exchanged by t_i and t_j, denoted by $V(c_{i \rightarrow j})$.

Definition 2: An Architecture Graph AG $= G(P, L)$ is an directed graph, where each vertex $p_i \in P$ represents one IP core in the architecture. Each directed edge $l_{i,j} \in L$ is a physical link connecting p_i to p_j.

Definition 3: The mapping of TG onto the processors of AG is defined by one-to-one mapping function $map(\mathcal{T})$:

$$\begin{aligned} map : \mathcal{T} \rightarrow P \quad \Longrightarrow \quad & p_s = map(t_i) \\ & \forall\ (t_i, t_j) \in \mathcal{T}^2 \ \ \exists\ (p_s, p_m) \in P^2 \mid p_s \neq p_m \end{aligned}$$

The constraint above guarantees that each task in TG can be assigned to an IP core. Figure 3.4 illustrates the execution trace of a given application task graph which has been mapped to an ONoC platform. t_i, t_j and t_k are assigned to p_s, p_m and p_r, respectively. We assume that each IP core is of the same type, so that the processing time of task is not relevant to the assigned IP.

Table 3.1: Application and Time Model Parameters.

Parameters	Description
t_i	Task i
$c_{i\to j}$	Communication from task t_i to task t_j
$V(c_{i\to j})$	Communication volume exchanged by t_i and t_j
N_t	Number of tasks in TG
N_c	Number of communications in TG
p_i	Processor i
$l_{i,j}$	Directed physical link connecting p_i to p_j
τ_p^i	Processing time of task t_i on one processor
τ_{end}^i	Completion time of task t_i
$pre(t_i)$	Set of tasks predecessors of task t_i
$T_{i,k}$	Set of occupied time slots by $c_{i\to k}$
$Pl_{i,k}$	Set of directed links traversed by $c_{i\to k}$
$W_{i,k}$	Set of reserved wavelengths for $c_{i\to k}$
B	Bit rate per wavelength

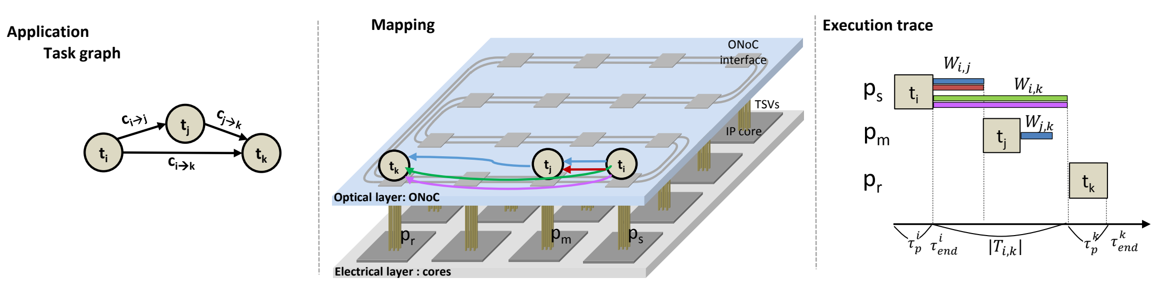


Figure 3.4: Execution trace of a given application task graph which has been mapped to an ONoC platform.

Table 3.1 summarizes the parameters involved in the model and equation 3.2 gives the communication time between task t_i and t_k, expressed as $|T_{i,k}|$ denoting the cardinality of set $T_{i,k}$.

$$|T_{i,k}| = \left\lceil \frac{V(c_{i\to k})}{|W_{i,k}| \times B} \right\rceil \tag{3.2}$$

$|W_{i,k}|$ denotes the cardinality of set $W_{i,k}$, i.e. the number of reserved wavelengths. The global execution time problem can be formulated as:

$$\begin{aligned} &Minimize \quad && Total\ execution\ time\ = Max\left(\tau_{end}^{k}\right)\ \forall\ k \in [0, N_t - 1] \\ & && \tau_{end}^{k} = \tau_{p}^{k} + Max\left(\tau_{end}^{i} + |T_{i,k}|\right)_{\forall\ i\ |\ t_i \in pre(t_k)} \\ &\text{with} && \forall\,(c_{i\to k}, c_{i\to j}) \in C \quad W_{i,k} \cap W_{i,j} = \varnothing \quad \text{if} \quad Pl_{i,k} \cap Pl_{i,j} \neq \varnothing\ \wedge\ T_{i,k} \cap T_{i,j} \neq \varnothing \end{aligned} \tag{3.3}$$

The constraint ensures that no wavelength is used more than once on a same waveguide segment at a given time.

As we know, multiple wavelengths channel is suitable for reducing communication time to reach an efficient network. Here, we define conflicting wavelengths when they propagate on a same segment of the waveguide (spatio conflict) at a given time (temporal conflict). As previously mentioned, conflicting wavelengths can interfere with each other through optical switching element, creating crosstalk noise. In order to analyse the conflict degree for an application running on ONoC, the conflict graph is introduced:

Definition 4: The conflict graph CG $= G(C, E)$ is an undirected graph, with each vertex $c_{i\to j} \in C$ represents the communication between t_i and t_j, and each edge defines the conflict relationship between two communications.

For example, the edge between $c_{i\to k}$ and $c_{i\to j}$ exists if temporal and spatial conflict for these two communications appear simultaneously. Namely, $c_{i\to k}$ and $c_{i\to j}$ share a same segment of waveguide concurrently. Therefore, $W_{i,k}$ and $W_{i,j}$ should not have wavelengths in common. In general, the more edges in the CG, the higher crosstalk introduced in ONoC.

3.3.3 Energy Model

The optimization process relies on on-chip laser since it impacts the performance of optical communications. We take the efficiency, the maximum output power, and the number of power levels available of laser into account.

The optical signals are emitted using on-chip VCSELs with direct modulation. The optical power depends on the laser modulation current, as illustrated in figure 3.5. In the OFF state, there is no current and hence no light emission. If the laser is likely to be used for data transmission, the current is set to i_{bias}, which corresponds to the threshold current from which light can be emitted. Since a direct modulation scheme is used, the data to be transmitted modulate the laser current as follows: data 0 leads to i_{bias} (no light emission) and data 1 leads to a higher current (light emission). In this work, we assume a driver allowing to configure the laser current for the transmission of data 1. The design complexity of the driver depends on the number of configurable laser power levels (NP_{lvl}). $NP_{lvl} = 4$ allows for selecting the modulation current for the transmission of data 1 as follow: i_1, i_2, i_3 and i_4 generate optical signals at $\frac{P_{max}}{4}$, $\frac{P_{max}}{2}$, $\frac{3P_{max}}{4}$, and P_{max} respectively. The maximum output power (P_{max}) is a designer choice regarding the targeted laser technology and the architecture requirements. In our simulations, we define it by considering the worst case regarding losses and communication configurations. $NP_{lvl} = 2$ leads to a reduced driver complexity but only allows for emitting optical signals at $\frac{P_{max}}{2}$ and P_{max}. $NP_{lvl} = 1$ is the baseline scenario, with only P_{max} for emitting data 1. Obviously, the higher NP_{lvl}, the more the ONoC is adaptable to application requirements (BER) and communication context (crosstalk and propagation losses). However, this adaptability comes with a higher design complexity.

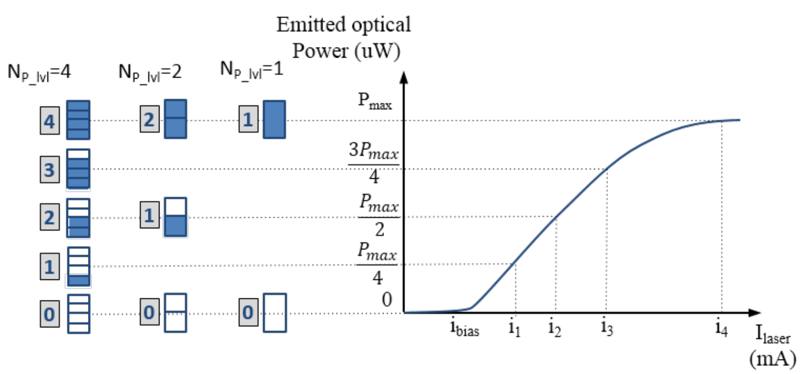


Figure 3.5: Configurable laser output power in the transmitter.

Table 3.2 summarizes the technological parameters and equation 3.4 calculates the ONoC laser energy consumption.

Table 3.2: Technological Parameters.

Parameters	Description
i_{bias}	Threshold current
P$_{max}$	Maximum output laser power
NP_{lvl}	Number of configurable laser power levels
$P_{laser,m}$	Optical power consumption for laser$_m$
η	Laser efficiency

$$\begin{aligned} Minimize \quad Total\ laser\ energy \quad &= \sum_{\forall\ c_{i\to k}\in C} \ \sum_{\forall\ \lambda_m \in W_{i,k}} \frac{P_{laser,m}}{\eta} \times |T_{i,k}| \qquad (3.4) \\ & P_{laser,m} = \frac{nP_{max}}{NP_{lvl}} \quad n \in [0, NP_{lvl}] \\ \text{with} \quad & BER_{\lambda_m} \geq target\ BER \end{aligned}$$

BER_{λ_m} can be evaluated by the BER estimation model introduced in chapter 2. The constraint shows that enough optical power ($P_{laser,m}$) is needed to meet the BER requirement.

3.4 Resolution

3.4.1 Offline Multi-Objective Optimization Algorithm

As previously mentioned, a trade-off between execution time and energy efficiency must be found. The objective problem is NP-hard and, in this case, heuristic approaches must be defined. Evolutionary technique, such as genetic algorithm, is one of the most popular choice [79]. It exploits a fixed-sized population of candidate solutions (called chromosome) and evolves over a number of generations to reach a satisfactory solution [80]. Our approach is based on a binary mixed integer coded genetic algorithm [81].

For example, a chromosome is defined as illustrated in figure 3.6. Since we reuse configuration mode Cfg_3 (figure 3.1) as an example, there are three interfaces, four wavelengths (i.e. four lasers per interface) and four laser output power levels. Tasks t_0, t_1, and t_2 are mapped on processor p_0, p_1, and p_2, respectively, which leads to optical communications between p_0 and p_1 ($c_{0\to1}$), p_0 and p_2 ($c_{0\to2}$), p_1 and p_2 ($c_{1\to2}$).

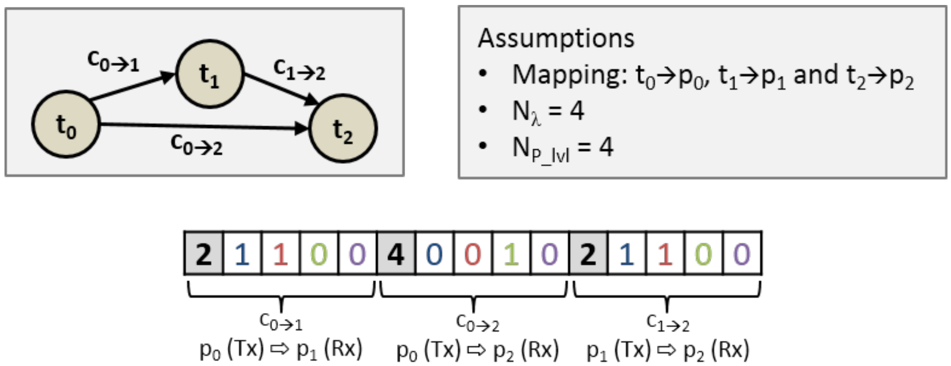


Figure 3.6: An example of chromosome.

It is divided into as many parts as there are communications. Each chromosome part consists of one integer element in grey box and N_λ binary element in white box. The first integer gene of each chromosome part gives the selected laser output power level: it corresponds to the laser configuration. The following N_λ binary genes correspond to the wavelengths utilization: value 0 or 1 indicates that a wavelength is used or unused, respectively. For the sake of simplicity, we assume that the allocated wavelengths of a communication use the same output laser power. Thus one integer value is sufficient for each chromosome part. In this example, both $c_{0\to1}$ and $c_{1\to2}$ use λ_0 and λ_1 on lasers power level 2. While $c_{0\to2}$ is implemented using only λ_2 on laser power level 4. Obviously, the more wavelengths are allocated for a given communication, the higher the bandwidth.

For the multi-objective optimization problem, the process works as follows. Figure 3.7 illustrates the design flow of the optimization. Firstly, a random parent population $\mathcal{A}$ is generated. Constraints validate the chromosome by ensuring that i) BER requirement is reached and ii) no wavelength is used more than once on a same waveguide segment. In case the constraints are not satisfied, this solution is considered invalid and directly set the fitness value of two objectives to infinity. For each valid solution generated, it is evaluated by estimations of i) the application execution time, considering the ONoC allocated bandwidth, and ii) the ONoC power consumption, considering the selected output power of the laser sources. Then we can extract the solutions which ensure the trade-off between execution time and energy efficiency. Afterwards, classical crossover and mutation operators are applied to evolve the initial population.

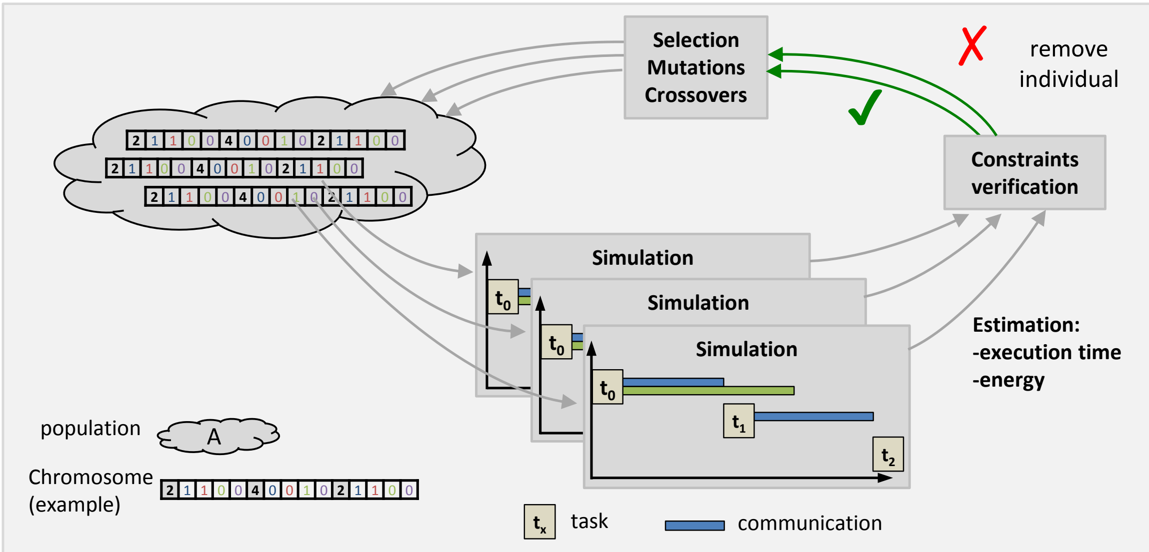


Figure 3.7: Offline multi-objective optimization.

Figure 3.8 explains how to generate new population by two basics genetic operators. These genetic operators are applied into each element of the chromosome separately, i.e. on each integer and binary elements. Figure 3.8 gives an example of modifying binary elements.

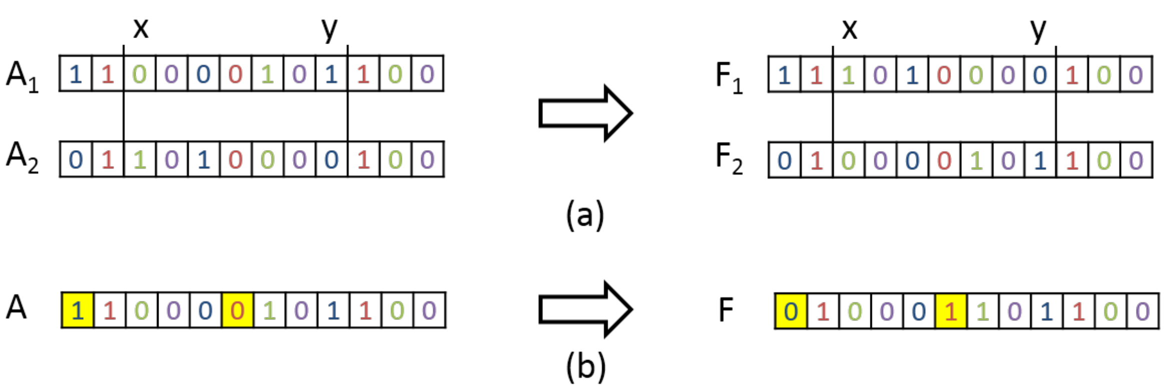


Figure 3.8: Genetic operators: a) two points crossover; b) swap mutation.

- Crossover: For two randomly chosen chromosomes (A_1, A_2), two random crossover points x, y are selected. The two offsprings are generated by exchanging the genes

$[x, y]$ of two parents.

$$\begin{aligned} \forall A_1, A_2 \in \mathcal{A} \quad F_1 &= A_1[:x[+A_2[x:y] + A_1]y:] \\ F_2 &= A_2[:x[+A_1[x:y] + A_2]y:] \end{aligned} \tag{3.5}$$

- Mutation: Two randomly genes are chosen and swapped into the chromosome.

$$A[\ldots, x_i, \ldots, x_j, \ldots] \underset{mutation}{\Longrightarrow} A[\ldots, x_j, \ldots, x_i, \ldots] \tag{3.6}$$

Finally a set of configuration modes on Pareto-front can be reported for high-level decision when a set of communications must be managed. Furthermore, the resulting configuration modes can be embedded in the controller and deployed at run-time according to execution requirement (e.g. high performance and low power), which is out of the scope of this thesis.

3.4.2 Complexity of Genetic Algorithm

In a Genetic Algorithm (GA), all possible chromosomes of a given encoding constitute the search space. As such, the maximum number of possible solutions that a GA must search is simply 2^N for the binary array chromosome, where N is the number of bits in a chromosome. As the chromosome length increases by 1 bit, the search space doubles. In our case, if N_λ increases by 1, the chromosome length increases by N_c bits and the search space increases by 2^{N_c}. Clearly, the length of a chromosome is important. As an example, imagine the GA will converge after examining 1/10 of the search space. Chromosome length increases by N_c bits, the GA-complexity of the problem will be $(1/10) \times 2^{N_c}$. This is clearly an exponential convergence time [82].

For our case, the length of chromosome depends on N_c and N_λ as detailed in section 3.2. Two metrics are considered in our evaluation: energy consumption and application execution time. In order to apply our algorithm on a complex task graph without increasing GA-complexity, one idea is to reduce the size of chromosome. Thus we propose to decompose the complex task graph into N_s small Sub Task Graphs

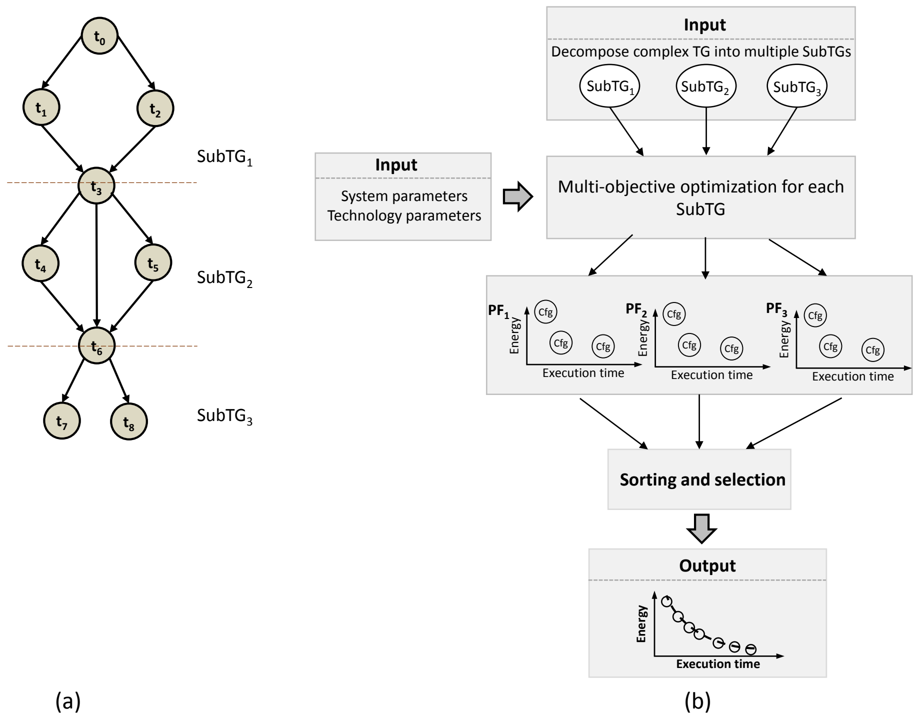


Figure 3.9: (a)Decomposition and (b) design space exploration of complex task graph.

(SubTGs) as shown in figure 3.9(a). Here we give a simple example, a 9-task graph is divided into 3 small SubTGs where each SubTG has only one root node. Firstly, for each SubTG$_k$ (with $k \in [1, N_s]$, N_s equals to the number of SubTGs), we follow the design flow detailed in section 3.4 to generate its Pareto front (PF$_k$) in terms of execution time and laser energy. Then for the complex task graph, the total execution time equals to the sum of all the execution time used by SubTGs and the energy consumption equals to the sum of all the energy used by SubTGs. After sorting and selecting, a global Pareto-front can be reported for high-level decision and the configurations for optimization of complex TG could be embedded in the system.

Obviously, we can not apply our methodology into all kinds of applications due to the complexity of chromosome. Moreover, we target the applications for which a significant amount of data are transmitted for each communication, i.e. the latency and dynamic power overhead used for the reconfiguration are negligible. We acknowledge that the approach we propose is not suitable for short-messages based data transmission such as those involved in cache-coherency protocols.

3.5 Optical Network Interface Management

3.5.1 ONoC Interface Configuration

Figure 3.10 illustrates a part of corresponding ONoC configuration of chromosome detailed in figure 3.6. The configuration of each interface is obtained as follows. First, optical channels are open by switching to the ON state the MRs involved in communications (i.e. MRs localized in the transmitter Tx of the source processor and the receiver Rx of the destination). Then, the power of the optical signals propagating through the channels is defined according to the selected lasers power level. In the example, at the transmitter of p_0 interface, three MRs are turned ON to implement communications $c_{0\to1}$ (λ_0 and λ_1) and $c_{0\to2}$ (λ_2). On the receiver side, the MRs corresponding to λ_0 and λ_1 are set to the ON state in p_1 interface, while the MR corresponding to λ_2 remains OFF to let the signal at λ_2 reaching p_2 interface where it will be dropped. In the chromosome part dedicated to $c_{0\to1}$, the laser output power level is set to 2: in p_0 interface, the lasers emitting at wavelength λ_0 and λ_1 are set to 50% of the maximum power. In the same interface, laser at λ_1 is set to 100% to match value 4 in corresponding gene for $c_{0\to2}$.

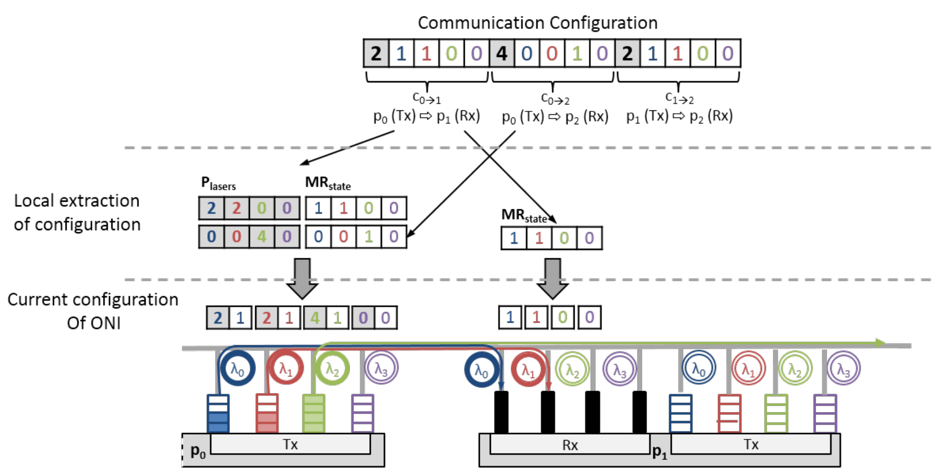


Figure 3.10: The corresponding ONoC configuration.

The size of an interface configuration is rather small since it requires only $N_\lambda \times (\log_2(N_{laser}) + 2)$ bits. Hence, it is easy and rapid to load predefined configuration mode from a dedicated memory in order to execute new applications or to adapt to an evolution of the execution context. Changing the ONoC configuration without loss of data can be simply achieved by waiting for the end of ongoing transmissions and by possibly setting the optical devices to the OFF state. However, lower latency reconfiguration and tighter adaptation to the evolution could be carried out, for instance, to maintain existing optical channels. This could be carried out at run-time by an operating system which is out of the scope of this thesis

This example illustrates the flexibility of our approach, which combines bandwidth allocation and laser output power tuning features. Indeed, we can allocate high bandwidth channels to time critical communications or minimize bandwidth for lower constrained communications. Tuning the laser output power ideally complements bandwidth allocation flexibility by i) alleviating crosstalk effect and by ii) allowing energy proportional optical communications.

3.5.2 Configuration Management

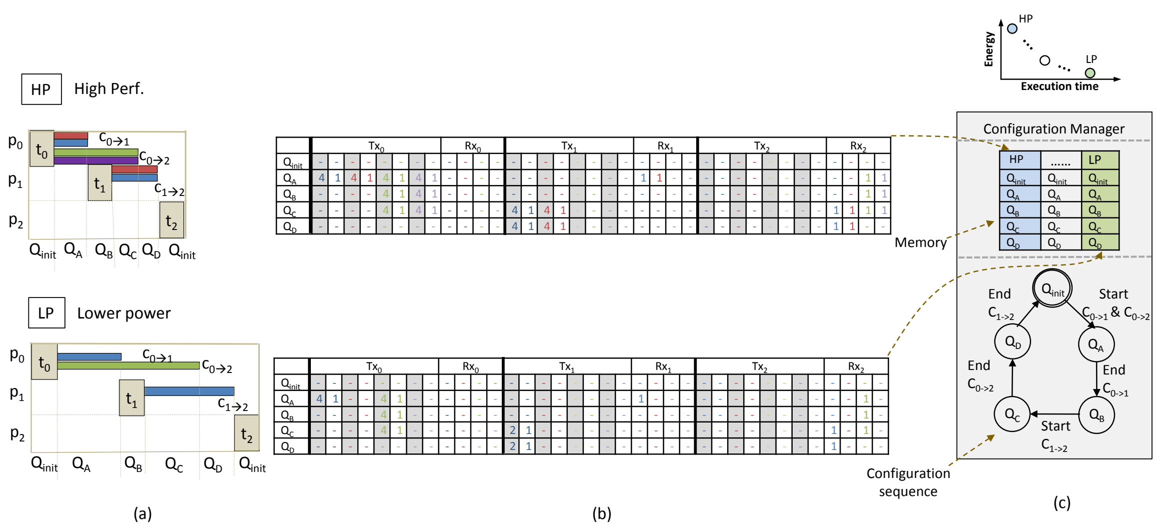


Figure 3.11: (a) Execution trace (b) related configuration memory of high performance and low power mode (c) configuration manager.

When a task graph enters our proposed framework, the offline simulation is carried out to generate performance and energy efficiency trade-offs and then the resulting solutions can be embedded in the system. Figure 3.11(a) shows the execution traces of two resulting solutions: high performance and low power of the task graph described in section 3.1. Then figure3.11(b) illustrates the related configurations of the transceivers (Tx_i) and the receivers (Rx_i) placed along waveguide (only values different from 0 are represented for clarity purpose). Q_{init} describes that all the optical devices are in the OFF-state, i.e. there is no communication in the ONoC. The table depicts sequential evolution of optical devices states during the execution of application. As previously stated, grey box indicates the laser level and white box corresponds to MR state. For example, four non-zeros values in the second line of grey box of HP represent four on-chip lasers emitting at wavelengths λ_0, λ_1, λ_2 and λ_3 with the maximum laser output power (here NP_{lvl}=4). Figure3.11 (c) illustrates the configuration manager where the configurations are stored in a dedicated memory. The configuration manager is responsible for opening optical channels when communications are initiated and for closing them when data transfers terminate. In order to satisfy specific requirement of system, we can download related configuration mode in the memory at run-time. Then the manager updates state of optical devices by following a configuration sequence. Each state of the configuration sequence is associated to a configuration Q.

As shown in figure 3.12, the suitability of configuration mode depends on the execution context (e.g. energy and time constraints). Before deploying a new configuration mode from dedicated memory, the specific constraint should be verified. Silicon photonics is a promising option for interconnects within datacenters, by using optical fiber as data transmission medium [83]. Here, we assume the example application executed sequentially in a battery powered system, at first, the abundant energy supply can drive the application runs as fast as possible. In this context, the manager takes use of high performance mode in the first iteration as shown in figure 3.12 . After 3 iterations of execution, a majority of energy is consumed so that the energy-saving mode should be considered. The manager thus loads Low Power mode for the last two iterations. In low power mode, p_0 sends a control signal to the configuration manager when t_0 execution ends and when data involved in $c_{0\to1}$ and $c_{0\to2}$ are ready for transfer. The manager switches from configuration Q_{init} to Q_A: Tx_0 and Rx_1 are reconfigured with the values stored in the configuration memory. Two on-chip lasers are powered on to emit optical signal for purpose of reducing energy consumption (the laser output powers are not illustrated for the sake of clarity) as shown in figure 3.12.(a). For low-latency purpose, the ONoC is reconfigured in parallel, which is implemented using dedicated electrical wires connecting the controller to the optical devices. Once the new

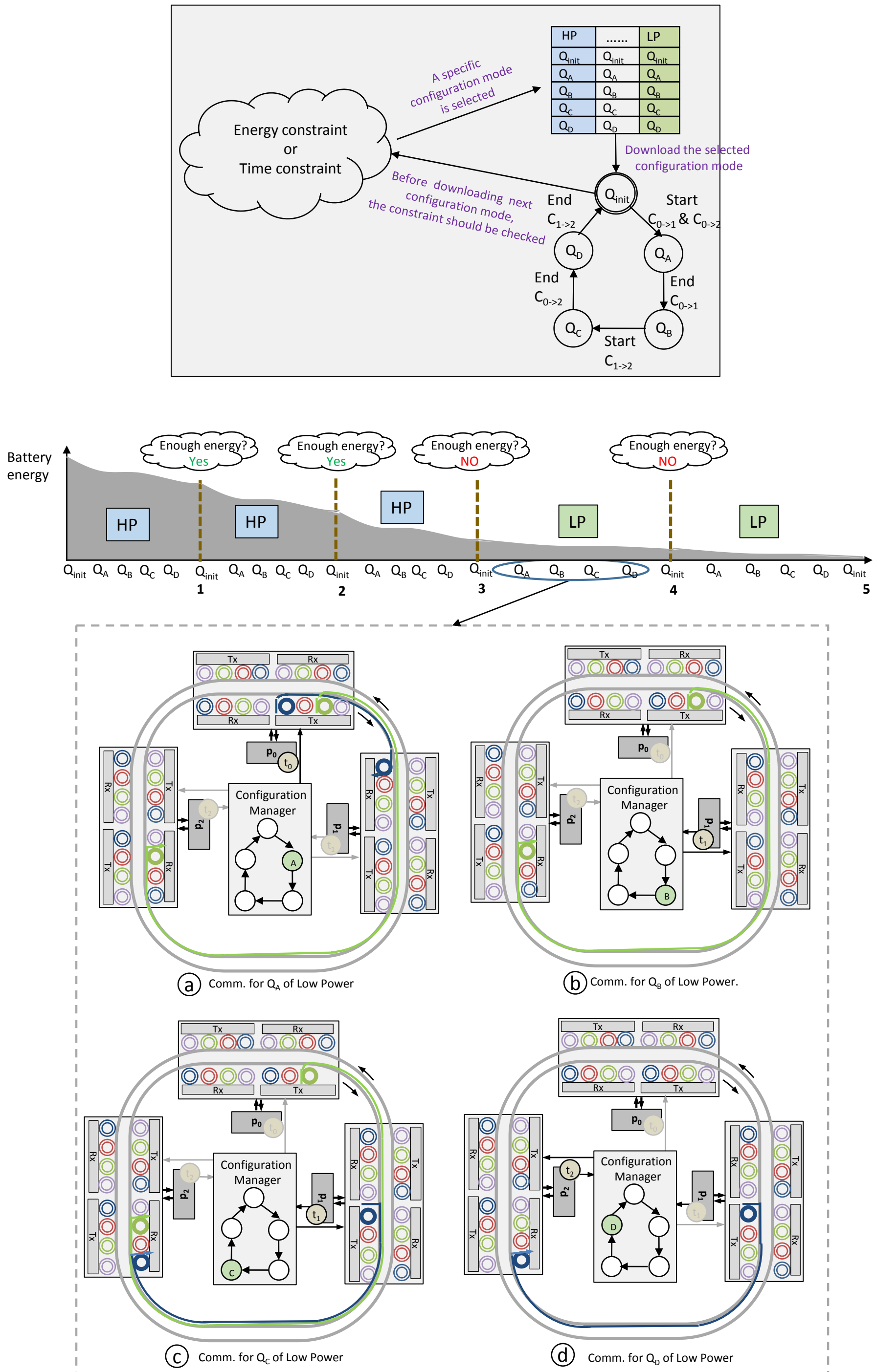


Figure 3.12: Reconfiguration of the ONoC at run-time.

configuration is loaded, the data transfer starts and, once it ends, another control signal is sent to the configuration manager to disable the communication channels. For this purpose, a new configuration (Q_B as shown in figure 3.12.(b) is loaded i.e. the lasers and the MR involved (blue one) in the communication are turned OFF. Figure 3.12.(c) illustrates communication $c_{1\to2}$ is carried out by the controller. The reconfiguration of the interfaces to implement a new communication has no impact on the ongoing communications: in this example, green signal on p_0 to p_2 waveguide segment is used for $c_{0\to2}$ while $c_{1\to2}$ is carried out using blue wavelength (reconfiguration $Q_B \to Q_C$). Then $c_{0\to2}$ ends and the green laser and MR are turned off as illustrated in figure 3.12.(d) (Q_D). Finally, the resources are released and become available for the next iteration.

3.6 Conclusion

Nanophotonic interconnects are promising solutions for high-performance and low-power on-chip communications. However, configuring an ONoC is challenging since the optimal use of optical resources depends on contradictory objectives related to BER requirements, expected execution performance and power budget. In order to get the best performance/energy trade-off, we combine, for the first time, wavelength allocation and laser power scaling for ring-based WDM ONoC. We introduce an offline framework based on genetic algorithm to get a set of Pareto-optimal solutions. Each solution considers application-dependant transmission between source and destination, which is critical for performance optimization. The framework handles several inputs, such as application, architecture, and device parameters. On the one hand, the exploration of design space at both device level and system level is allowed; on the other hand, a set of performance and energy efficiency trade-offs can be generated once the input parameters are defined. Then the designer can select the most appropriate solutions at run-time according to the specific requirement. Evaluations are carried out for a 3D WDM-based ring optical interconnect and the simulation results are shown in the chapter 4.

Chapter 4

Optical NoC Architectural Exploration

An offline framework has been proposed in chapter 3 to find the trade-off of performance and energy efficiency for ONoC-based MPSoC. It also allows the exploration of design space at both device level and system level. Key input parameters at device level including lasers, photodetectors sensitivity, waveguide losses and MR transmission are taken into account. At system level, application mapped on an ONoC (e.g. task execution time, data size and BER requirement), architectures (e.g. number of wavelengths, number of waveguides, number of interface and distance between interfaces) are considered. Based on a set of device level and system level input parameters, established analytical models (e.g. BER model, energy model, and time model) allow the evaluation of energy efficiency and performance of the considered optical interconnects. Finally the resulting ONoC configuration modes on a Pareto front from the multi-objectives optimization can be obtained. Our framework is generic and can work with different device and system inputs. In this chapter, we first conduct evaluations of design space in terms of various input parameters under a given application in section 4.1, then the trade-off of performance and energy efficiency based on defined input parameters under some complex applications is presented in section 4.2.

4.1 Case Study

To illustrate our approach, we study in this section a given application presented in figure 4.1. Regarding the architecture, we assume a Chameleon interconnect composed of 4×4 ONIs with 1GHz processor frequency and 1Gb/sec laser data rate. The task computation times and the volume of transmitted data are given in kilo clock cycles (kcc) and kilobit (kb) respectively. The mapping we assumed is illustrated on the right hand side of the figure. The wavelength allocation and power distribution are combined to investigate the impact of different ONoC design parameters on performance and power consumption. The multi-objective optimization algorithm is set to iterate 800 times with a population size of 500 chromosomes. Some technological parameters are summarized in table 5.6.

Table 4.1: Technological parameters.

Parameter	Value	Ref
Waveguide propagation loss	-0.274dB/cm	[37]
Photodetector sensitivity	-20dBm	[84]
Laser efficiency	15%	[84]

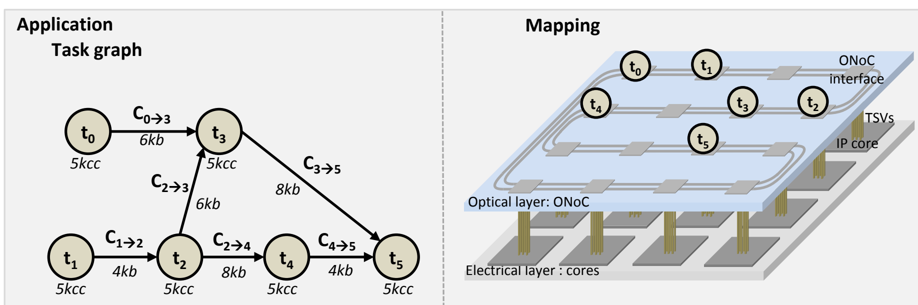


Figure 4.1: Characterized task graph and its mapping on a 4×4 ONoC.

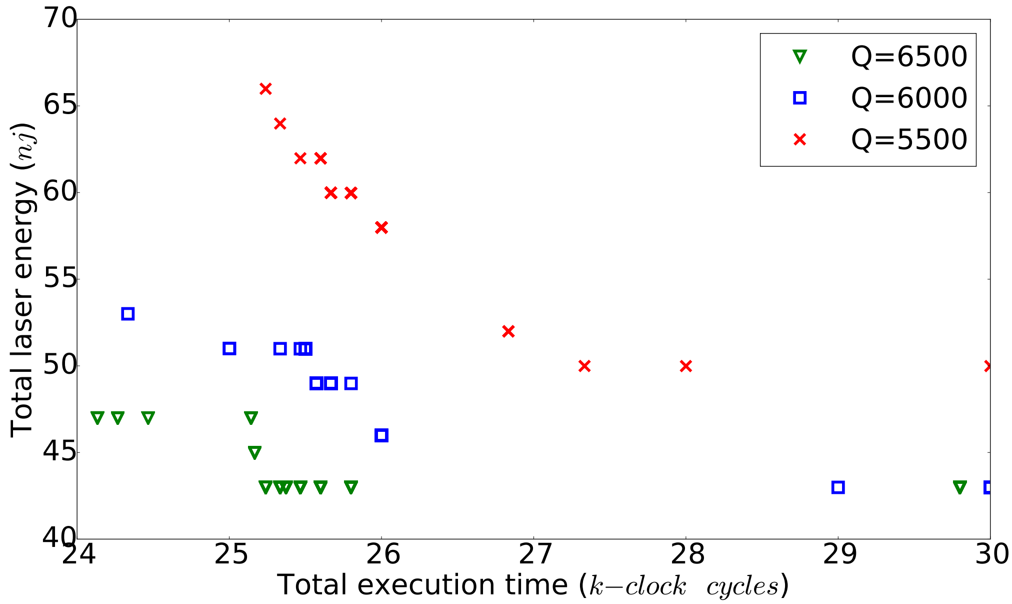
4.1.1 MR Characteristic Parameters

As mentioned in MR transmission model (see section 2.2.3), FSR and Q have a great impact on the BER and energy model for WDM-based ONoCs. Understand the energy variations based on employing different values of theses parameters, we analyse the energy consumption in a given application while considering different values of parameters. It is worth mentioning that the values of Q and FSR can be tuned by varying MR diameters. We assume 8 wavelengths per waveguide and the number of laser level $NP_{lvl} = 7$. As we consider electrical power lasers from 1mW to 4mW, the configurable powers, in mW, are [1, 1.5, 2, 2.5, 3, 3.5, 4].

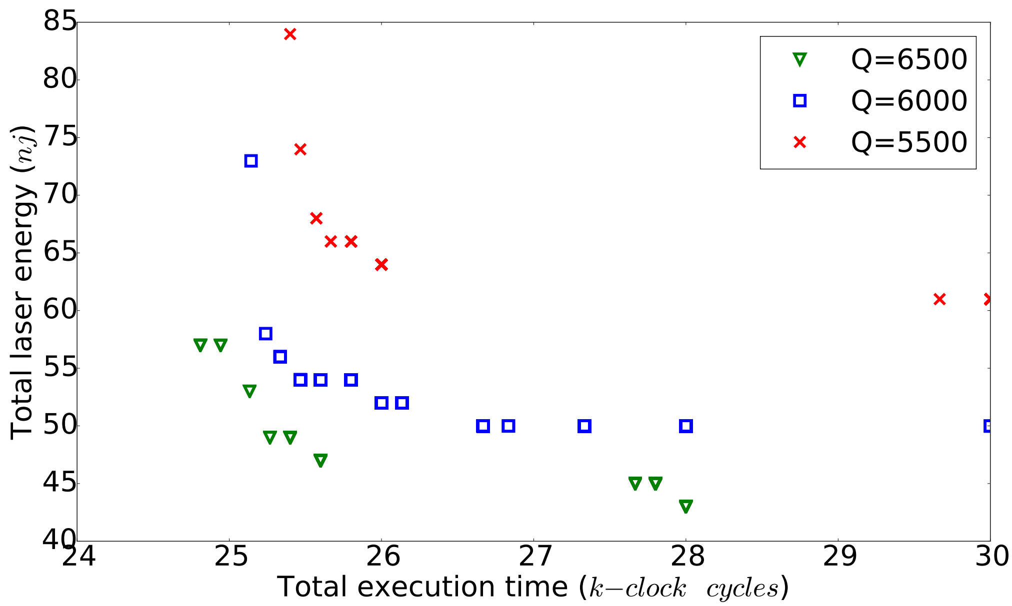
4.1.1.1 Q-factor

In this study, we investigate the impact of Q-factor on the ONoC energy consumption. For this purpose, we assume FSR equals to 8nm and Q varies between 5500, 6000 and 6500. Figures 4.2(a) and 4.2(b) show the resulting solutions for targeted BER of 10^{-9} and 10^{-12}, respectively. Each solution is characterized by the total laser energy consumption and the application execution time. The total laser energy is the sum of all the energy used by the lasers. These results highlight the Pareto front induced by the conflicting optimization of performance and low-power objectives. As shown in figure 4.2(a), the higher the number of signals propagating simultaneously, the lower the execution time and the higher the laser output power needed. For instance, when Q equals to 5500, 30kcc execution time leads to 50nj energy consumption while 25.3kcc execution time leads to 66nj energy consumption. As Q increases, the full width at half-maximum bandwidth decreases, which results in a lower inter-channel crosstalk and a lower propagation loss along optical path, and consequently smaller energy consumption. For instance, for the solution characterized by 25.8kcc execution time, when $Q = 5500$, the energy consumption is 59nj, compared to 48nj and 42.5nj for $Q = 6000$ and $Q = 6500$ respectively. Additionally, the energy consumption reduces more deeply when Q is small. Moreover, the time efficiency is improved with Q. Precisely, Q varies from 5500 to 6500, the minimal execution time drops from 25.2kcc to 24.1kcc. This is due to higher quality MRs along the waveguide can support more wavelengths for a target BER. However, the higher Q, the higher circumference of MR, which means that less MRs can be fit into a given area, therefore providing lower integration density [85].

Figure 4.4(b) shows the same trends for a 10^{-12} BER. This set of solutions can be used if more robust executions of the application are needed, but this comes with a higher energy consumption since the laser output power needs to be increased to improve the SNR. For instance, for the solutions characterized by 26kcc execution time and with $Q = 5500$, the required energy increases from 58nj to 64nj when a BER of 10^{-12} is targeted. This trend illustrates the opportunity introduced by our approach to consider Quality of Service in ONoC architectures.



(a)



(b)

Figure 4.2: Optimization results: solutions are characterized by an execution time and an energy with Q=5500, 6000, and 6500 for a targeted BER of (a) 10^{-9} and (b) 10^{-12}.

4.1.1.2 FSR

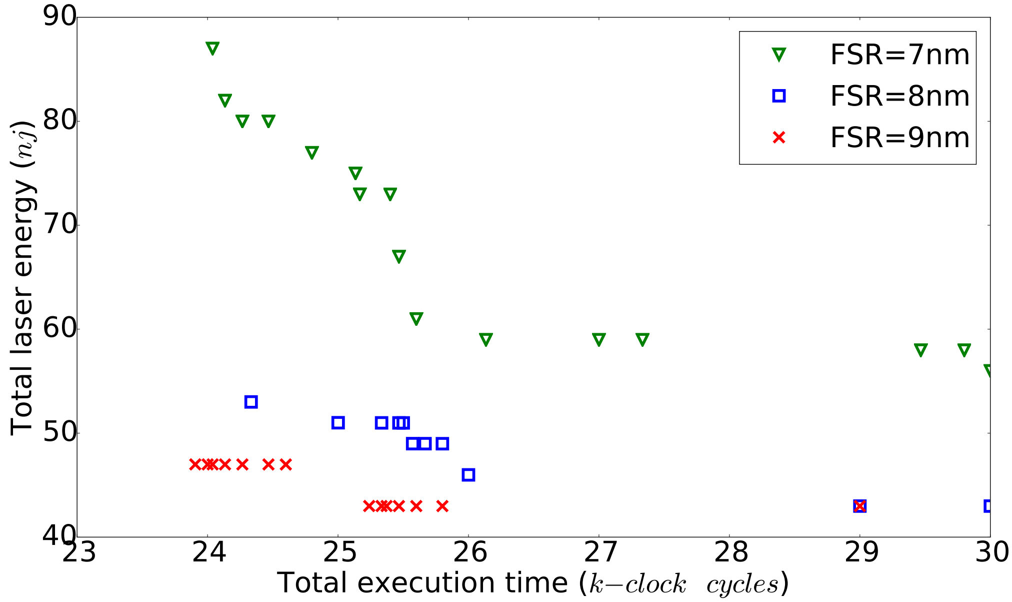
Figure 4.3 (a) and (b) show the execution time and energy consumption variation for a targeted BER of 10^{-9} and 10^{-12} respectively. We consider the use of $Q = 5500$ and FSR varying among 7, 8 and 9nm. It is worth noticing that, a much lower energy consumption is obtained due to larger FSR. For instance, when the execution time is 25.5 kcc, $FSR = 7nm$ leads to 65nj energy consumption, compared to 51nj and 42nj energy for $FSR = 8nm$ and $FSR = 9nm$, respectively. Indeed, as FSR increases, the spacing among different wavelengths is larger, which results in a smaller inter-channel crosstalk noise and insertion loss of MR, and consequently smaller energy consumption. Moreover, the higher FSR, the smaller circumference of MR, which can provide higher integration density. Depending on this context, a small MR is expected to construct WDM-based ONoC. Then the same trend is obtained for a 10^{-12} BER as shown in figure 4.3(b).

To summarize, a smaller MR leading to a higher FSR can fit in more wavelengths and have a higher aggregated data bandwidth. However, Q is proportional to the circumference of MR. In order to achieve low-power objective, a proper small size of MR without sacrificing the Q-factor should be considered.

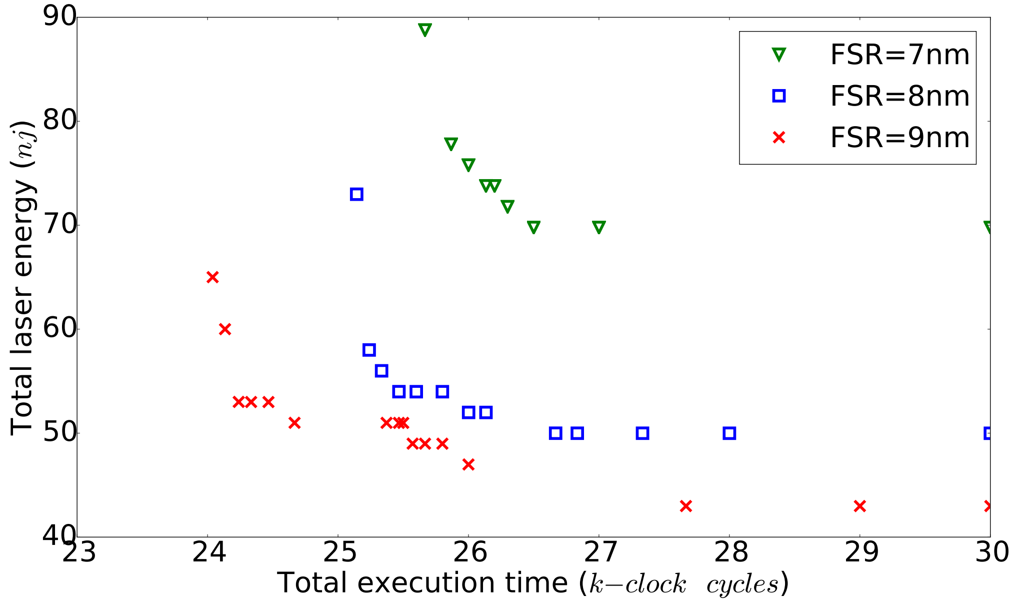
4.1.2 Number of Laser Power Levels

In this study, we investigate the impact of the number of laser power levels on the ONoC energy consumption. For this purpose, it is assumed that the Q is 6000 and FSR is 8nm. We assume a single waveguide with $N_\lambda = 8$ and we assume three values for NP_{lvl}: 3, 5 and 7. As we consider electrical power lasers from 1mW to 4mW, the configurable powers, in mW, are: [1, 2.5, 4] for $NP_{lvl} = 3$, [1, 1.75, 2.5, 3.25, 4] for $NP_{lvl} = 5$, and [1, 1.5, 2, 2.5, 3, 3.5, 4] for $NP_{lvl} = 7$.

Figures 4.4(a) and 4.4(b) show the resulting solutions for targeted BER of 10^{-9} and 10^{-12}, respectively. In Figure 4.4-a, $NP_{lvl} = 3$ leads to solutions ranging from 56nj to 63nj for the laser energy and from 26kcc to 24.2kcc for the execution time. These results highlight the Pareto front induced by the conflicting optimization of performance and low-power objectives. This trend is further increased for $NP_{lvl} = 5$ since the execution time interval ranges from 30kcc to 24kcc. It is important to notice that the energy consumption decreases since the optical signal power emitted by the laser can be more



(a)



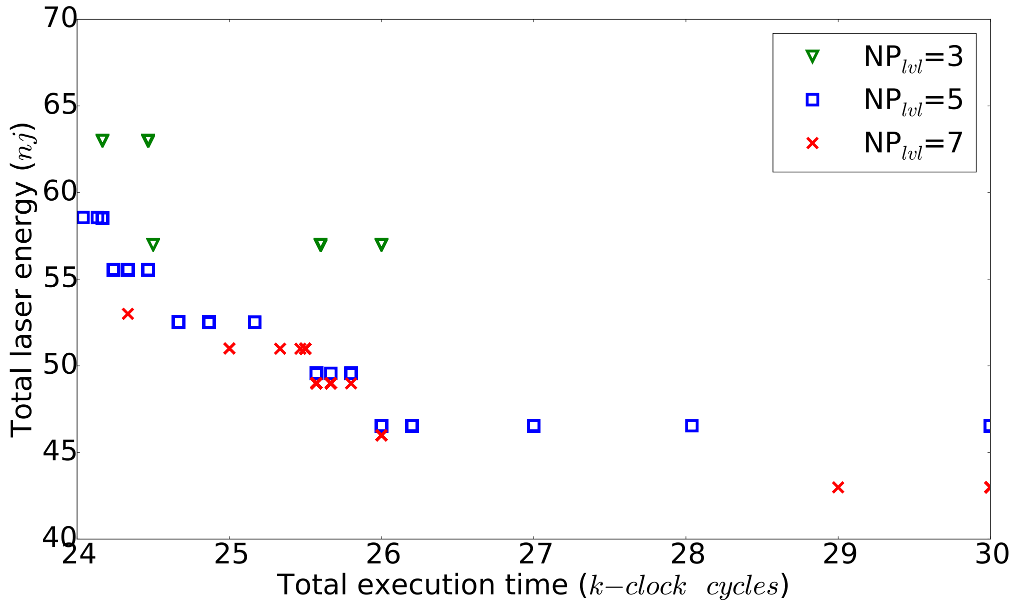
(b)

Figure 4.3: Optimization results: solutions are characterized by an execution time and an energy with FSR=7,8, and 9 for a targeted BER of (a) 10^{-9} and (b) 10^{-12}.

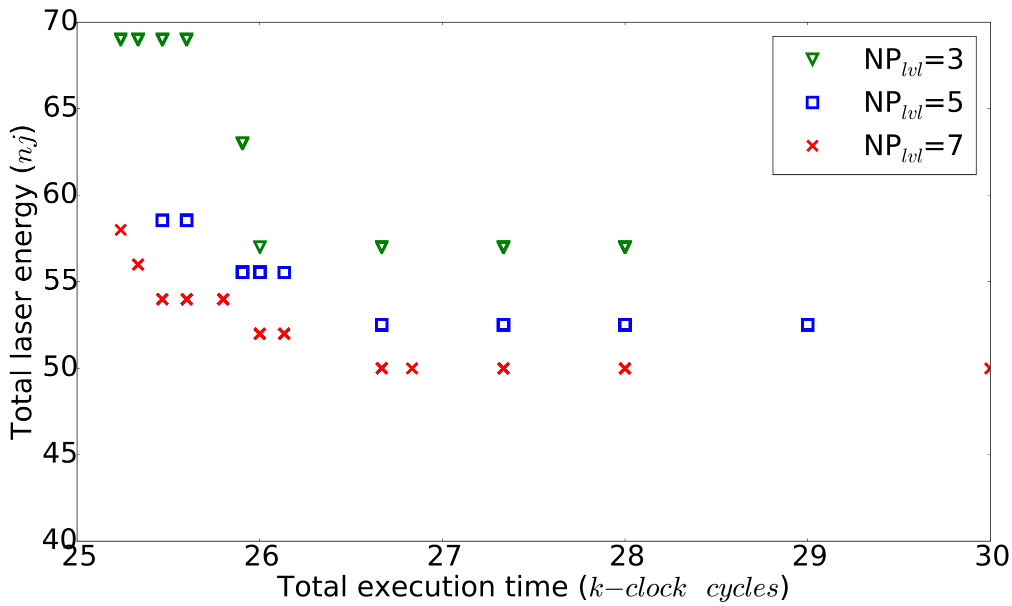
accurately adapted to compensate the losses experienced by the optical signals and to reach the BER requirement. For instance, for the solutions characterized by 26kcc execution time, $NP_{lvl} = 3$ and $NP_{lvl} = 5$ lead to 56nj and 46nj energy consumption, respectively. The energy is further reduced to 45nj for $NP_{lvl} = 7$, which allows for finer-grain configuration of the laser. Besides, the energy gain of $NP_{lvl} = 3$ to $NP_{lvl} = 5$ is greater than that of $NP_{lvl} = 5$ to $NP_{lvl} = 7$. Obviously, the higher NP_{lvl}, the closer to the optimal solution the results. However, this comes with a more complex design of the laser driver, which is out of the scope of this work. These results demonstrate the need for configurable laser output power to improve both performance and energy consumption in ONoC based architectures. Figure 4.4(b) shows the same trends for a 10^{-12} BER. A higher energy is consumed since the laser output power needs to be increased to improve the SNR. For instance, for the solutions characterized by 26kcc execution time and with $NP_{lvl} = 7$, the required energy increases from 45nj to 51nj when a BER of 10^{-12} is targeted.

4.1.3 Number of Wavelengths

Figures 4.5(a) and 4.5(b) give the optimization results for $N_\lambda = 4$, 6, and 8 and with $NP_{lvl} = 7$ and $NP_{lvl} = 5$, respectively. The targeted BER is 10^{-12}. The results first highlight that the higher the number of wavelengths, the higher the number of solutions and the higher the diversity. For instance, for $N_\lambda = 4$, the execution time of resulting solutions ranges from 30kcc to 28.3kcc, while for $N_\lambda = 8$ solutions ranges from 30kcc to 25.3kcc. However, more wavelengths in the ONoC leads to a reduction of the energy efficiency. Indeed, more wavelengths in the ONoC means more MRs in the ONIs, which increases the losses. The energy efficiency reduction is further accentuated when multiple wavelengths are assigned to implement the communications: solutions showing low execution time lead to inter-channel crosstalk effects that require higher laser output power. For instance in Figure 4.5(a), the configuration providing the lowest energy consumption with $N_\lambda = 8$ requires 50nj and 30kcc, while the lowest energy consumption with $N_\lambda = 6$ requires only 43nj for the same execution time. The same execution time is synonym of the same number of allocated wavelength for each communication, meaning the same amount of dynamic power is used. Hence, the difference in terms of energy comes from the static losses due to the increase of MR

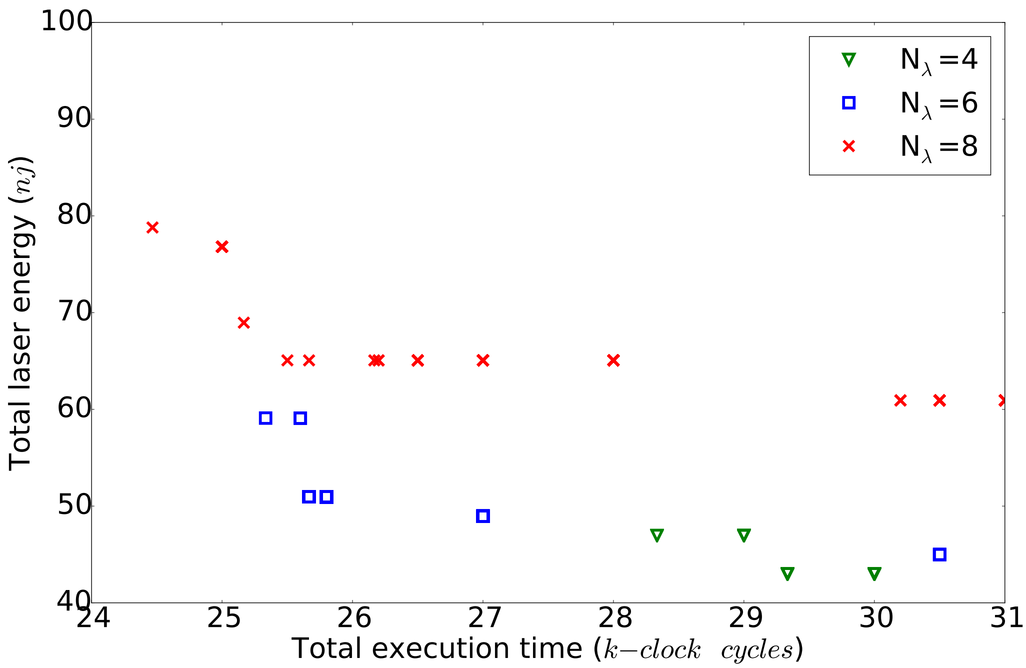


(a)

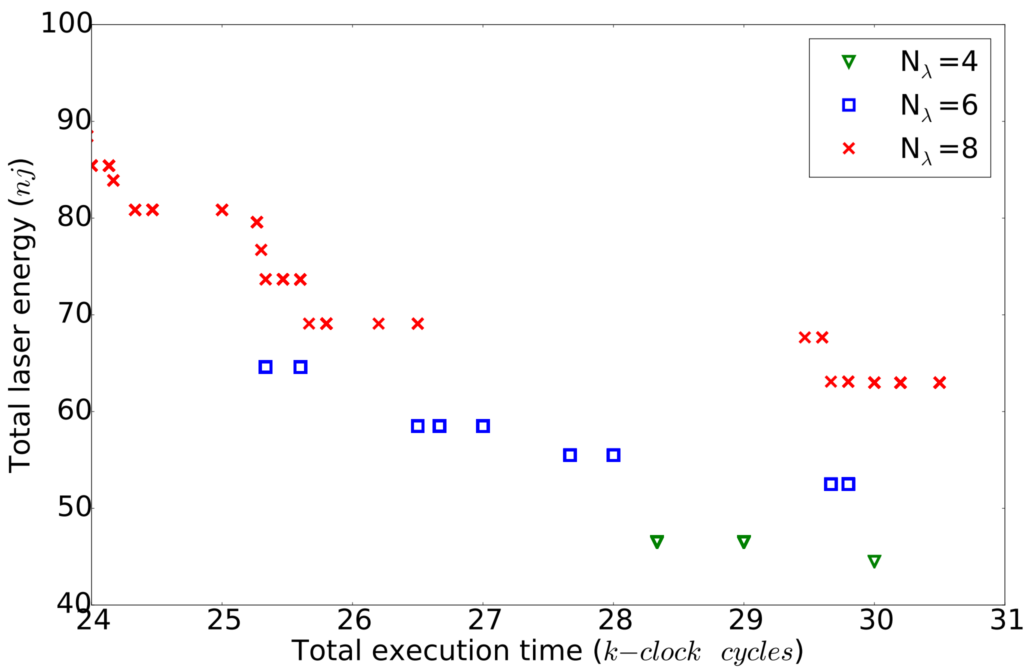


(b)

Figure 4.4: Optimization results: solutions are characterized by an execution time and an energy for NP_{lvl}=3, 5, and 7 for a targeted BER of (a) 10^{-9} and (b) 10^{-12}.



(a)



(b)

Figure 4.5: Optimization results for various number of wavelengths N_λ, and with (a) NP_{lvl}=7 and (b) NP_{lvl}=5.

numbers.

To summarize, designing an ONoC with more wavelengths increases the bandwidth allocated for each communication. However, this also reduces the energy efficiency due to higher propagation losses (more resonators are crossed) and higher crosstalk (the wavelength spacing is reduced).

In order to illustrate the complexity of our methodology, table 4.2 lists the number of valid solutions generated by the genetic algorithm and the number of solutions located on the Pareto front. For instance, for 8 wavelengths, 64290 valid solutions have been generated. Among them, only 14 solutions provide an efficient trade-off between energy and execution time. To illustrate theses results, figure 4.6 draws all the solution on a energy consumption versus execution time plot for 8 wavelengths. We can note that a huge number of solutions are far away from the Pareto front. This figure confirms that the exploration of performance and energy consumption trade-off is critical for efficient communication.

Table 4.2: Amount of generated valid solutions and amount of solutions on the Pareto front.

Number of WL	Number of solutions on Pareto front	Number of valid solutions
4	4	34389
6	7	41459
8	14	64290

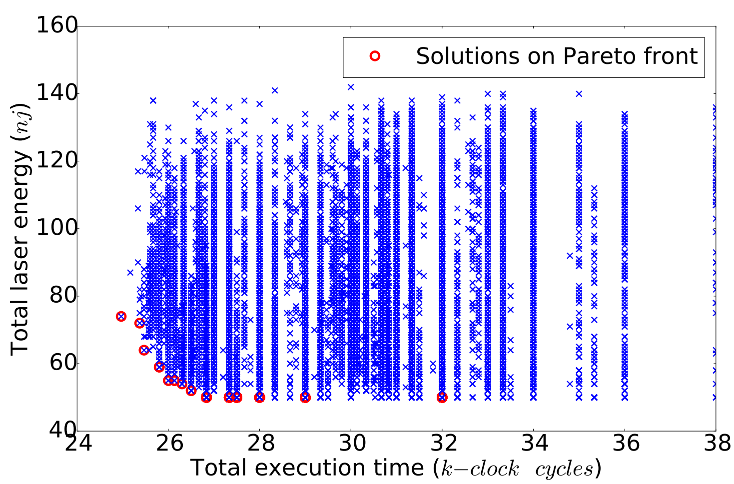


Figure 4.6: Valid allocation solutions generated for 8 wavelengths.

4.1.4 Task Mapping

An application, represented by a task graph, should be assigned to a set of available cores by mapping different functions to different regions of the system. A good mapping solution based on specific communication patterns is one of best chance to improve energy efficiency. On the top of figure 4.7, 3 different mapping solutions are illustrated. Mapping0 is the case we have studied above. Based on an application mapped on an

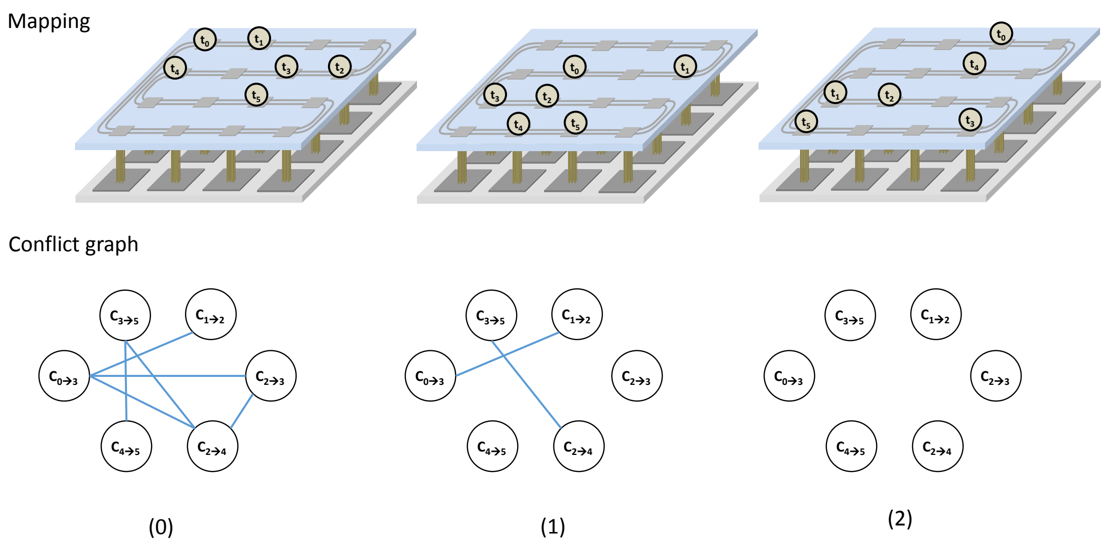
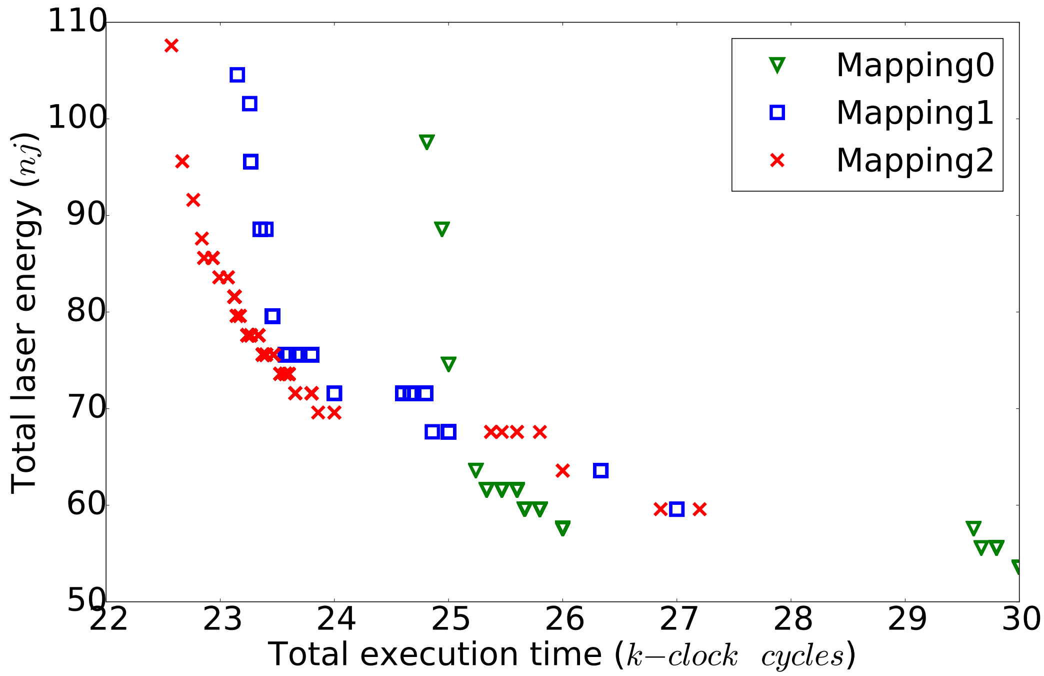
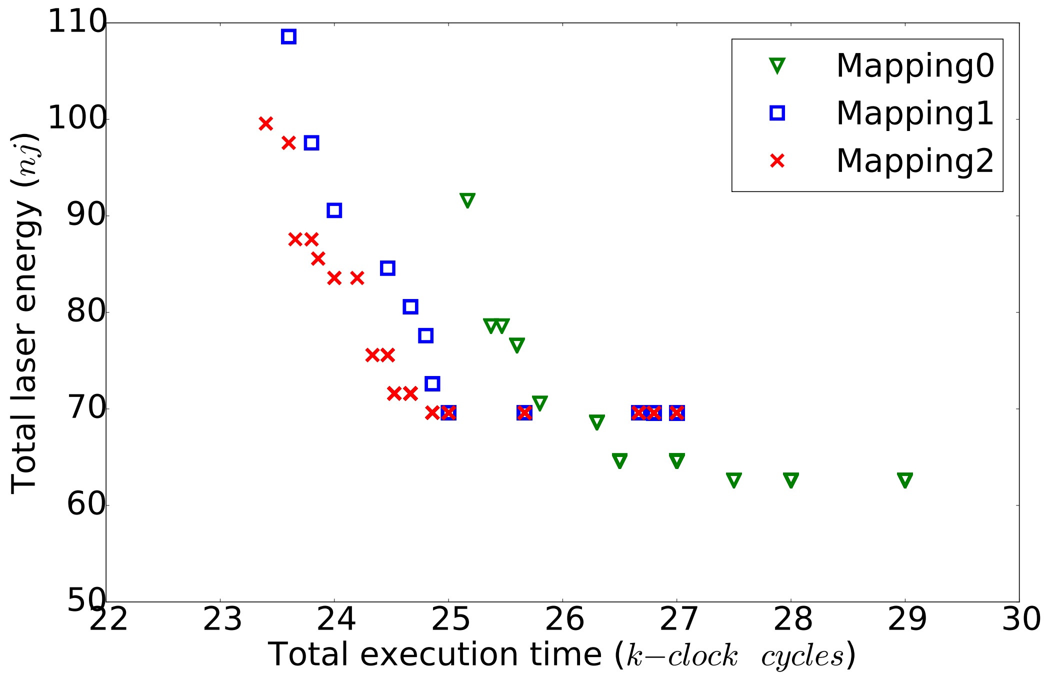


Figure 4.7: Three mapping solutions with their related conflict graphs representing the worst case before wavelength allocation. Each edge appears when spatial conflict occurs and temporal conflict is not clear only with task dependency relation.

ONoC, the spatial conflicts between different communications are clear. However, due to lack of wavelength allocation, detailed temporal conflicts between communications are not clear. Fortunately, task dependency relation gives some ideas in the study of temporal conflict. For example, $c_{0\to3}$ and $c_{3\to5}$ should be executed successively, thus no temporal conflict between them and consequently no edge between them in the CG. On the bottom of figure 4.7, it illustrates the related CG representing the worst case before wavelength allocation. Zero edge exists in the CG of Mapping2 due to the absence of spatial conflict. They are avoided by the use of bidirectional transmission and waveguide partitioning properties of considered architecture. On the whole, the CG gives a global vision of conflict when an application running on ONoC. The more edges in the CG, the higher crosstalk introduced in the network.



(a)



(b)

Figure 4.8: Optimization results for various mapping with target BER of (a) 10^{-9} and (b) 10^{-12}.

In this part, we investigate the impact of task mapping on the ONoC energy consumption and execution time. Each waveguide integrates 8 λ. Q and FSR are set to 5500 and 8nm respectively. The configurable powers are [1, 1.5, 2, 2.5, 3, 3.5, 4]mW.

Figures 4.8(a) and 4.8(b) give the optimization results for three different mappings with target BER of 10^{-9} and 10^{-12}, respectively. Insertion loss and crosstalk noise are two key factors that affect energy consumption of photonic interconnect. When the number of parallel wavelengths propagating in the waveguide is small, the inter-channel crosstalk is also small. In this case, the energy efficiency is primarily limited by the insertion losses accumulated along the optical path. To avoid inter-communication crosstalk, Mapping2 sacrifices the proximity of highly-communicating tasks, leading to high propagation loss and consequently high energy consumption. For instance, the solution characterized by 26kcc, Mapping2 consumes 62nj energy while Mapping0 consumes 58nj energy as shown in figure 4.8(a). However, the energy consumption of Mapping2 is exceeded by Mapping0 when the execution time is higher than 25kcc. Since inter-channel crosstalk becomes main effect of the reduction of energy efficiency. Moreover, a improvement of the minimal execution time is reached. Here 22.6kcc for Mapping2, compared to 23.2kcc for Mapping1 and 24.8kcc for Mapping0. This is due to higher bandwidth channel reached by Mapping2 with non-overlapping communications.

To conclude, task mapping allows the tasks assigned into different regions so that inter-communication crosstalk can be avoided. However, it risks assigning the high-communication tasks into distinct and separate cores, which results in high propagation losses along optical path. Mapping of application tasks to the core is thus tough, especially if low energy consumption is expected for a target BER.

4.2 Scalability of the Approach

In this section, we investigate the scalability of the approach using ONoC architectures with 16, 32, and 64 ONIs, that interconnect 64, 128, and 256 cores respectively. Each ONI is connected to a cluster of four electrically connected cores. We assume two waveguides, eight wavelengths per waveguide, and $NP_{lvl} = 5$. For the ONoC with 64 cores the possible electrical laser power values are [2, 4, 6, 8, 10] mW, and for 128 cores the possible electrical laser power values are [3, 6, 9, 12, 15] mW, while for 256 cores the laser power values are [4, 13, 22, 31, 40] mW. Regarding the applications, we use a random task graph generator that provides applications including from 52 to 107 tasks and from 80 to 158 communications. The task execution time values are randomly selected between [100, 1000] cc and the communication volumes are randomly selected within [100, 1000] bytes range. The targeted BER is 10^{-9} and each task is randomly mapped on a dedicated core. As we assume shared memory within a same cluster, no latency is assumed for intra-cluster communications.

	Graph ID	Number of Tasks	Number of Comm.	Energy (nj)					
				Low Power			High Perf.		
				ON-OFF	Proposed	Reduc. (%)	ON-OFF	Proposed	Reduc. (%)
64 cores 16 ONI	TG 1	55	80	463	141	69.5	463	187	59.6
	TG 2	52	78	433	118	72.7	433	194	55.2
	TG 3	57	82	469	102	78.3	469	120	74.4
	TG 4	60	92	525	141	73.1	525	195	62.9
	TG 5	63	93	497	128	74.2	497	193	61.2
	TG 6	62	92	530	147	72.3	530	236	55.4
	TG 7	56	87	480	103	78.5	480	148	69.2
	TG 8	63	91	492	112	77.2	492	156	68.3
	Average			**486**	**124**	**74.5**	**486**	**179**	**63.3**
128 cores 32 ONI	TG 9	107	158	1346	253	81.2	1346	389	71.1
	TG 10	94	139	1145	213	81.4	1145	314	72.6
256 cores 64 ONI	TG 9	107	158	2871.4	748.35	73.9	2871.4	1051	63.4
	TG 10	94	139	3055.2	654.4	78.6	3055.2	1153.4	62.2

Table 4.3: Energy comparisons between ON-OFF strategy ($NP_{lvl} = 1$) and our Proposed methodology with $NP_{lvl} = 5$ for different task graphs.

	Graph ID	Number of Tasks	Number of Comm.	Energy (nj)			Execution Time (kcc)		
				Low Power	High Perf.	Variation	Low Power	High Perf.	Variation
64 cores 16 ONI	TG 1	55	80	141	187	1.33	41.2	23.8	1.73
	TG 2	52	78	118	194	1.64	37.9	23.4	1.62
	TG 3	57	82	102	120	1.18	39.9	21.7	1.84
	TG 4	60	92	141	195	1.38	37.3	23	1.62
	TG 5	63	93	128	193	1.51	41.2	21.8	1.89
	TG 6	62	92	147	236	1.61	43.2	24.7	1.75
	TG 7	56	87	103	148	1.44	31.7	20	1.57
	TG 8	63	91	112	156	1.39	37.5	23.4	1.60
	Average			**124**	**179**	**1.44**	**38.7**	**22.7**	**1.71**
128 cores 32 ONI	TG 9	107	158	253	389	1.54	62.19	37.5	1.66
	TG 10	94	139	213	314	1.47	64.38	33.6	1.92
256 cores 64 ONI	TG 9	107	158	748.35	1051	1.40	64.38	38.1	1.69
	TG 10	94	139	654.4	1153.4	1.76	61.01	34.58	1.76

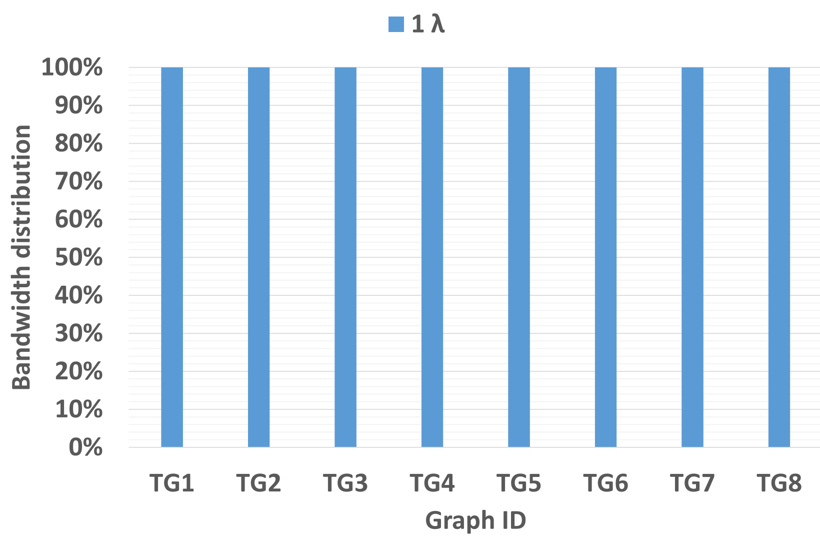
Table 4.4: Energy and execution time for Low Power and High Performance wavelength allocations strategies.

Tables 4.3 and 4.4 summarize the characteristics of the task graphs and the optimization results. Since our approach leads to a Pareto front, we only show the solutions with i) the lowest energy consumption (denoted Low Power in the table) and i) the lowest execution time (High Perf.). We also provide results for the baseline ON-OFF strategy (i.e. $NP_{lvl} = 1$). Table 4.3 compares the energy consumption between ON-OFF strategy and our proposed method. Regarding the low-power solutions, our methods leads to 74.5% reduction on average in the energy consumption compared to the ON-OFF solutions. This is due to the laser output power which is adapted to satisfy the BER requirements. Indeed, for our proposed strategy, low Laser power levels are often selected to support communications, while in the ON-OFF strategy, only high Laser power level can be used. It is worth mentioning that the high-performance solutions resulting from our method also shows a significant 63% reduction on average in energy compared to the baseline strategy. We can note that for both strategies, because the number of wavelengths allocated for each task graph is the same, the execution time is unchanged.

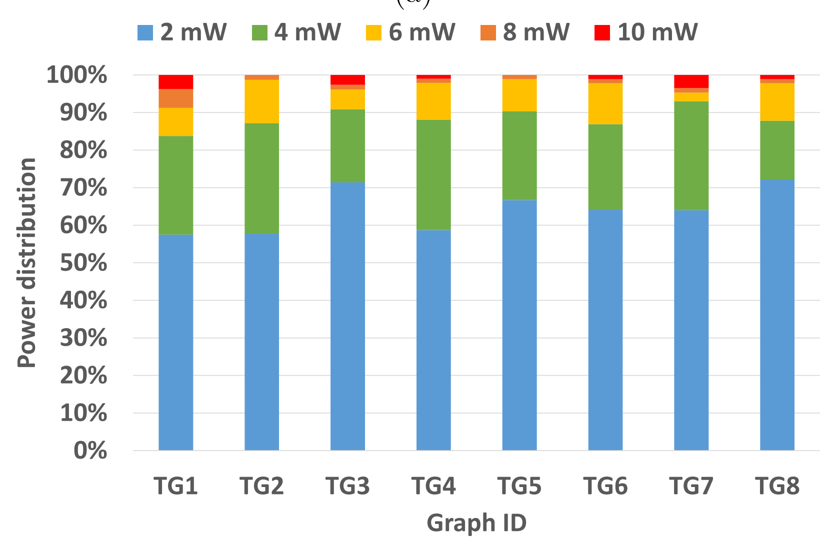
As shown in Table 4.4, the solutions offer, on average, 44% energy variation and 71% execution time variation trade-offs for 64 cores architectures. As the results show, the execution time gain increases for larger architectures, while the energy gain decreases. Indeed, for these results, we use random task placement which leads to long communications between tasks, and hence high Laser power to support these communications. This simulation context is the worst as possible, and we can imagine that a smart task placement should reduce these long communications and gives the opportunity to select more often the low Laser power levels. Nonetheless, these results demonstrate the efficiency of flexible wavelengths allocation and laser output power tuning to adapt the nanophotonic interconnects according to application requirements.

Figure 4.9(a) and 4.9(b) give the distribution of allocated bandwidth and laser power levels for the low-power solutions. All the communications are allocated on a single wavelength for a minimum crosstalk effect. On average, the 2mW power level is selected in 60% of the cases and is followed by 4mW (15% on average). Higher power levels are selected to compensate for the losses experienced by long range communications. Figure 4.9(c) illustrates results for high-performance solutions. Taking TG8 as an example, 13% of the communications are carried out using a single wavelength, while 20% rely on all the wavelengths. Figure 4.9(d) shows the laser power distribution for each task graph for the same solutions. 60% and 20% of the lasers are configured to 2mW and 4mW respectively for TG8. The lasers configured with a higher output power emit signals experiencing significant losses (i.e. long distance communications) or crosstalk. Hence, the resulting ONoC configuration modes do not simply allocate all the wavelengths for each communication, even though maximum execution performance is reached. Indeed, our methodology allows to compute the maximum wavelengths to be allocated by considering simultaneous communications and computation time that may not be reduced.

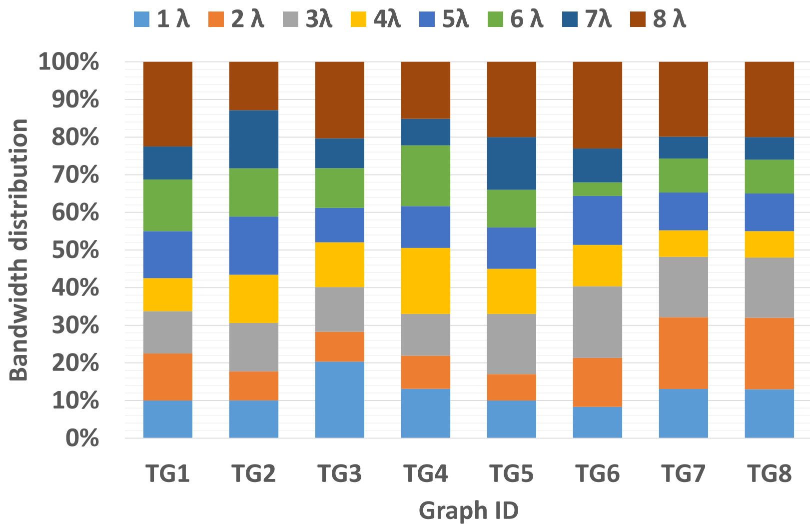
These results demonstrate that the proposed approach to combine laser output power and wavelength allocation reaches the maximum execution performance while saving energy by using only the required optical resources. Furthermore, complex applications lead to solutions with very different energy and execution time figures.



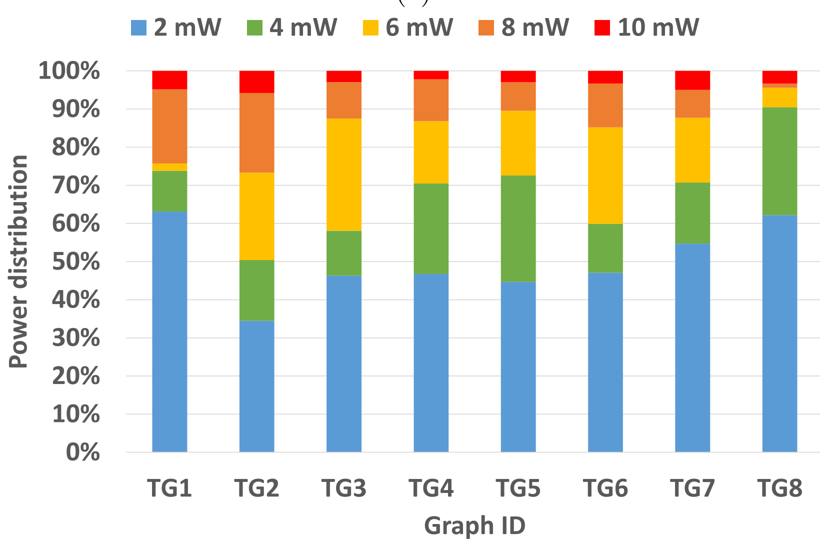
(a)



(b)



(c)



(d)

Figure 4.9: Distribution of allocated bandwidth and laser power levels for (a-b) low-power solutions and (c-d) high-performance solutions.

This confirms the opportunity induced by our offline approach to adapt the ONoC configuration mode at run-time according to the execution context and QoS requirements.

4.3 Conclusion

We have proposed an offline methodology allowing to find energy efficiency and performance trade-off including 1) low-power solutions, which tend to minimize the number of used wavelengths, and 2) high-performance solutions, for which multiple wavelengths are allocated to shorten the communication time. The methodology is formed as a framework handling several inputs, such as application, architecture, and device parameters. On the one hand, the exploration of design space at both device level and system level is allowed; on the other hand, a set of performance and energy efficiency trade-offs can be generated once the input parameters are defined. Hence, our proposed methodology is evaluated in two aspects. In the first step, the exploration of design space at both device level and system level is carried out under a given application. The evaluation on MR characteristics parameters (i.e Q-factor and FSR) shows that high Q-factor and large FSR of MR can introduce a small switching loss, which results in a low energy consumption. To further improve the energy efficiency, the configurable laser can provide just-enough power to meet the application BER. Then the N_λ is explored to provide high diversity of solutions. In the second step, the scalability of methodology is investigated with some complex applications under defined input parameters. As an example, for a 63-task application running on 64 cores, the relative variation in the execution time and energy is 60% and 39% respectively. Solutions showing good energy-performance trade-offs are also found. Compared to baseline solutions for which laser power is fixed, our method leads to 74% energy reduction on average.

Chapter 5

TbNoC: Electrical Tree-based Network for online management of ONoC

Offline wavelength allocation produces optimized solutions thanks to the global view of optical resources and complex algorithms used to explore design space of the solutions. However, it is restricted to predefined application and not able to handle a dynamic workload. In this case, dynamic policies to allocate the optical resources at runtime in a MPSoC architectures should be considered. In this chapter, we address this problem and propose an electrical tree-based network, TbNoC, to support dynamic management of optical resource. The proposed network can be combined to different ONoCs based on shared waveguide. For performance reason, we partition the global electrical control network into several hierarchical central-controlled agents and centralize the resource arbitration for purpose of reducing set-up path latency. Besides, our online wavelength allocation can be jointly used with offline approach for mixed deterministic and non-deterministic tasks running on ONoC. First, the wavelength allocation for deterministic tasks can be realized at compiled time by our proposed offline framework, which results in a set of offline configurations stored in a specific memory. Then, when a non-deterministic task occurs, the channel set-up process is launched to allocate wavelengths from those which are not used by offline configuration. To simplify the presentation of TbNoC, and to highlight its flexibility, we start this chapter by only considering online configuration; i.e. only non-deterministic tasks need to be supported. The capacity to support both offline and online strategy is then discussed at the end of this chapter. To illustrate the performance of TbNoC, we instantiate this control network to Chameleon ONoC, named Chameleon-TbNoC. As a result, we show that Chameleon-TbNoC improve a SNR of 1.6dB compared to SUOR under uniform traffic. It is important to notice that the distribution of laser power is not considered in our online approach. We assume a fixed laser power source which is defined by considering the worst case regarding losses of architecture. The online management of proportional laser power will be considered in our future work.

5.1 Key Concepts of Tree-based Control Network

5.1.1 Partition Principle

Definition: a tree topology T=(V,E) is a symmetric graph, where N_{lvl_0} is the number of clusters and N_{in} is the number of input ports of each agent. Then the number of tree level NT_{lvl} equals to $log_{N_{in}}(N_{lvl_0})$.

$$V = \left\{ \begin{array}{ll} C_j & |j \in [0, N_{lvl_0}] \\ A_{i-j} & |i \in [1, NT_{lvl}], j \in [0, \frac{N_{lvl_0}}{(N_{in})^i} - 1] \end{array} \right\}. \tag{5.1}$$

$$E = \left\{ l(i,j) | i \in [0, NT_{lvl} - 1], j \in [0, \frac{N_{lvl_0}}{(N_{in})^i} - 1] \right\} \tag{5.2}$$

For the sake of clarity, each edge $l(i,j)$ connects A_{i-j} and $A_{(i+1)-j}$, except $l(0,j)$ connects C_j and A_{1-j}. The edge is represented by a bidirectional arrow instead of two unidirectional arrows.

An example of TbNoC is depicted in figure 5.1(a). 4 cores in a core cluster (C) are connected to each other by electrical link (the cores are not represented for clarity purpose). The intra-cluster communication is locally managed in the cluster and inter-cluster communication is managed by TbNoC. Each cluster contains a local cluster router, connecting to TbNoC for long-distance channel set-up and to optical layer for long-distance data transfer. As illustrated in figure 5.1(b), the clusters are placed in the bottom of the tree. Each cluster is labelled with C$_j$ where j is the cluster ID (indexed from 0 to 63). A pair of coordinates (l-p) is used to label each agent, where l denotes a node's level and p denotes its position within that level. In this example, each leaf

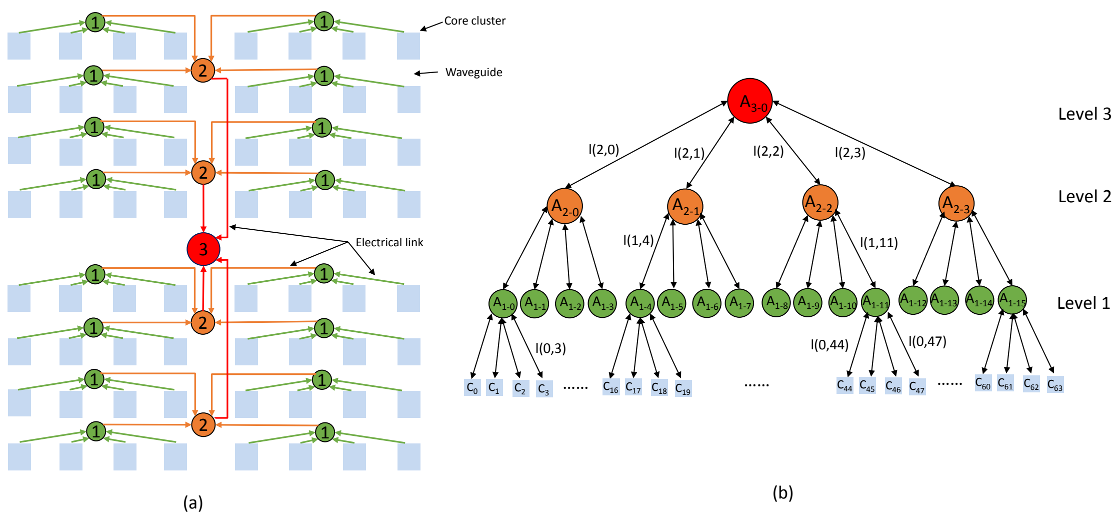


Figure 5.1: (a)Layout of 256-cores (b) 3-level TbNoC

level agent is connected to 4 clusters ($N_{in} = 4$), denoted by A_{1-j} (indexed from 0 to 15, green one). 4 consecutive leaf level agents are linked up to an unique branch level agent A_{2-j} (indexed from 0 to 3, orange one). Finally, the root agent A_{3-0} (red one) connects 4 branch level agents. Each control packet from the source cluster climbs the tree from low-level agent to high-level agent until reaching the first common ancestor of the source and destination core, then the wavelength allocation is activated to compute which wavelengths should be selected. Each level of agent can handle the channel set-up request. However, in our design, a specific memory for wavelength state information (wavelength table) is only employed in the leaf agent. Since in the case the tasks are well placed, a majority of communications will appear between neighbouring clusters and the related request will be handled in leaf agent.

5.1.2 Online Communication Protocol

Our protocol in TbNoC is designed to manage optical resources in ONoC. It can manage several existing ONoC, such as Chameleon and SUOR. When a core produces a request packet, the cluster router first check where the destination core is located.

Case 1: destination core is located on the same cluster as the source core does, the local fully connected electrical link is used for transmission and do not need to set-up an optical path.

Case 2: destination core is located on one neighboring cluster and the leaf agent is the first common ancestor of source and destination cores, the request will be handled in leaf agent.

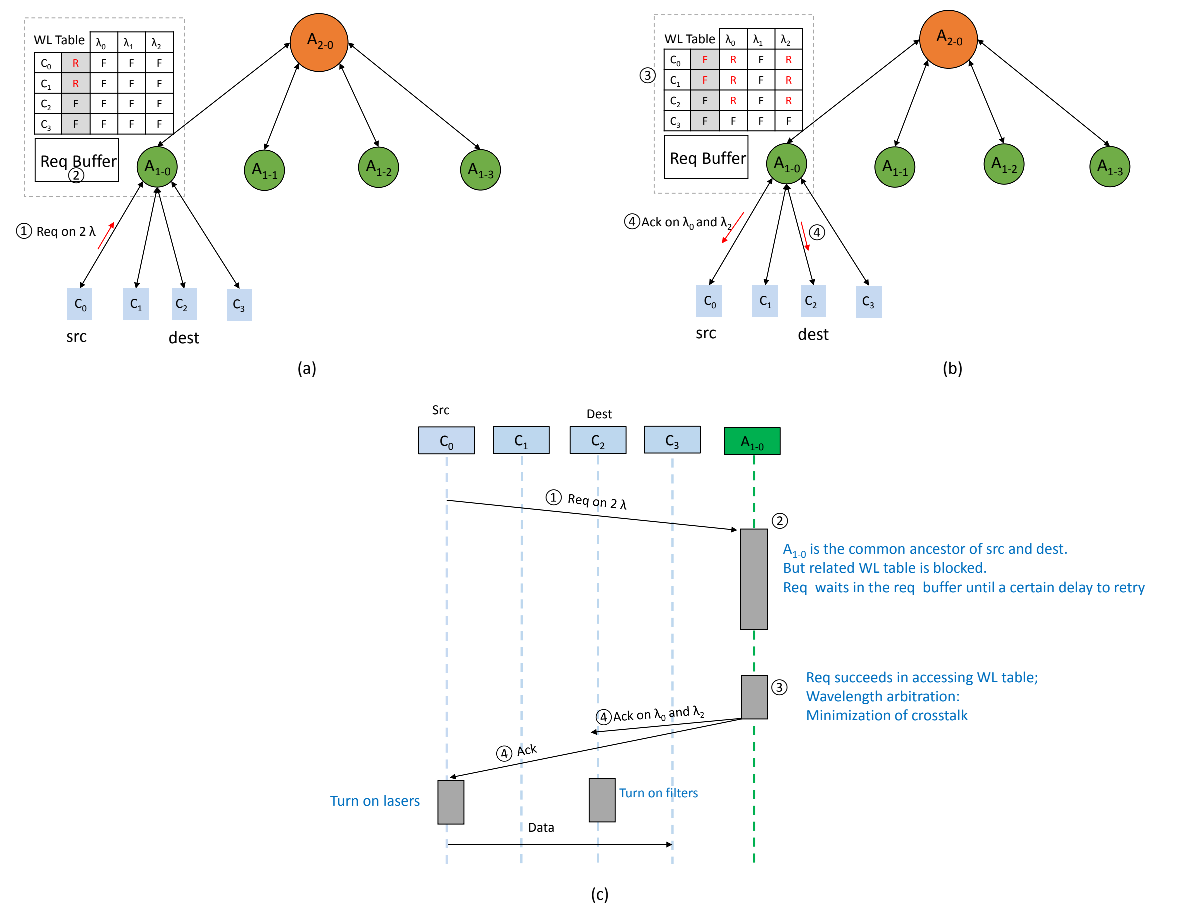


Figure 5.2: Wavelength reservation for an inter-cluster communication between a source cluster C_0 and a destination cluster C_2. (a) Request forward (b) Acknowledge success of optical path set-up (c) Protocol diagram.

The figure 5.2 shows us the process of wavelength allocation for an inter-cluster communication realized in A_{1-0}. The allocation process is initiated by a source cluster C_1 through the injection of a Req for requesting two wavelengths in the control network (mark ①). As we know, each leaf agent contains a wavelength (WL) state table (Only

WL table in A_{1-0} is appeared for sake of clarity). The size of wavelength table is $N_{in} \times N_\lambda$ if one waveguide is assumed in the ONoC. This table contains two parts: grey box represents the availability of wavelength table: Free (F) or Reserved (R). Only one request is allowed to access the same part of wavelength table at a time. White box represents the state of each wavelength : Free (F) or Reserved (R). As illustrated in the figure 5.2, we assume that C_0 and C_1 are not accessible, thus the Req should wait in a specific buffer (mark ②). After a certain latency, this Req retries to read the wavelength table. Two proper wavelengths are reserved according to common wavelengths state (mark ③). The common wavelength state is the size of N_λ bits and defined as below: when a wavelength is free along the whole communication path, this wavelength is set to Free, otherwise, it is set to Reserved. Finally, the Acks on λ_0 and λ_2 are sent backward to the source/destination cluster. Then the corresponding lasers/MR filters can be powered on and get preparation for sending/receiving the optical signal (mark ④).

Case 3: destination core is located in a distant cluster and the branch agent is the first common ancestor of source and destination core, the request will be handled in branch agent.

In Fig 5.3, C_0 initiates a communication with C_{14} on 2 wavelengths. Since A_{1-0} is not the common ancestor of source and destination clusters, the message is forwarded to A_{2-0} (mark ①). Due to the lack of wavelength state information, A_{2-0} thus broadcasts a message noted WL_STATE_Req to the related leaf agents (mark ②). Then WL_STATE_Req reserves the related rows of wavelength table in leaf agent. For example, $\{C_0,C_1,C_2$ and $C_3\}$ in A_{1-0} and $\{C_{12},C_{13},C_{14}\}$ in A_{1-3} are updated from F to R. Besides, the WL_STATE_Rps with related common WL_STATE is sent back to A_{2-0} (mark ③). Then the wavelength allocator calculates which wavelengths should be reserved for purpose of minimizing crosstalk and new generated Acks reserving λ_0 and λ_2 are sent back to the related leaf agents (mark ④). It also frees the related partition of wavelength tables and reserves the selected wavelengths ID. Finally, Acks are sent backward to C_1 and C_{14} for configuring MRs and lasers on the optical layer (mark ⑤).

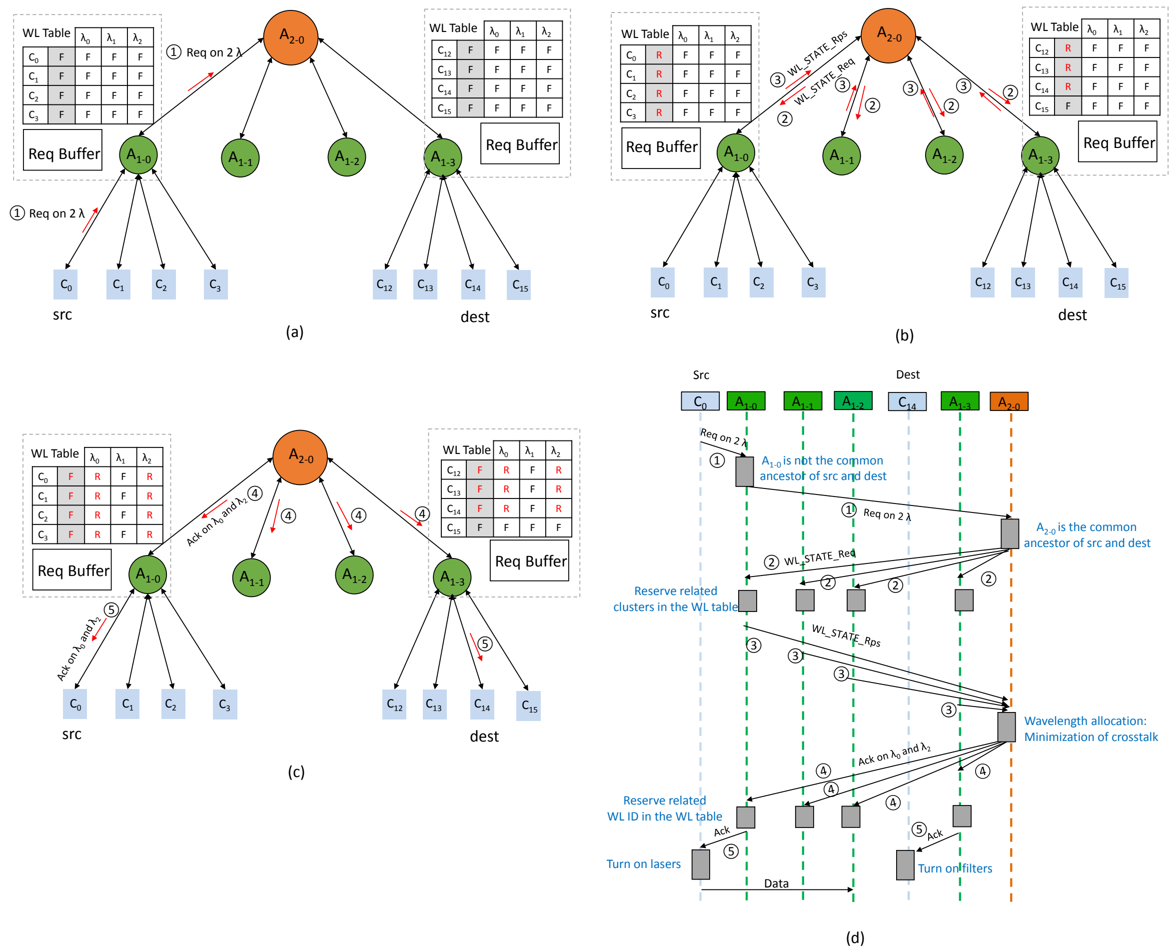


Figure 5.3: Wavelength reservation for an inter-cluster communication between C_0 and C_{14}. (a) request forward (b) exchange the wavelength state information (c) acknowledge the success of optical path set-up (d) protocol diagram.

5.1.3 Control Message Structure

Control packets are used to set-up optical paths for data transfer. Table 5.1 illustrates the structure of defined control messages. It contains information of the message type, the ID of the source IP (Src), the ID of the destination IP (Dest), the requested number of wavelengths (Req_NUM_WL) and wavelengths states (WL_STATE) .

Regarding Req_NUM_WL, the more the requested wavelengths, the higher the bandwidth between source and destination, the smaller the transmission time of data stream. When a request message is emitted, the requested number of wavelength

Type	Src	Dest	Req_NUM_WL	WL_STATE

Table 5.1: Representation of request message.

Message type	Use
Req	Try to establish an optical path
Ack	Validate the allocation of optical resources
Rel	Release the reserved optical channel
WL_STATE_Req	Try to read the wavelength state of related clusters
WL_STATE_Rps	Response of common wavelength state
WL_STATE_Rel	Cancel the reservation process and release the wavelength state table

Table 5.2: Use of different messages.

is determined (e.g., Req_NUM_WL is set to 2, meaning that two wavelength is requested). WL_STATE is the size of N_λ bits indicating whether a wavelength is Reserved (R) or Free (F) to realize the corresponding communication. Control messages can be divided into two parts according to its function. The use of messages are detailed in the table 5.2. First part contains Req, Ack, and Rel, which are used to set-up and release optical path. Second part contains WL_STATE_Req, WL_STATE_Rel and WL_STATE_Rps, which is responsible for exchanging wavelength state. Since the wavelength state table is only deployed in leaf level agent in our design, the arbitration handled in branch level agent should exchange the second part of messages with leaf agent to obtain corresponding wavelength state informations.

5.2 Implementation

5.2.1 Leaf Agent

Figure 5.4 depicts a part of example TbNoC where the micro-architecture of leaf agent is detailed. 4 clusters input ports are followed by 4 first-in-first-out (FIFO) queues for messages buffering. Since one packet in each time slot is allowed to deliver to the next functional module, Round-Robin (RR) algorithm is employed to decide which port has grant to access when a conflicting situation is found. Then another arbiter gives priority to the packets from A_{2-0} port to access control unit. In this way, the upper agent can read the wavelength states of related clusters as soon as possible. Once the control unit is reached, the algorithm 1 is executed. Each type of message is managed differently as detailed below:

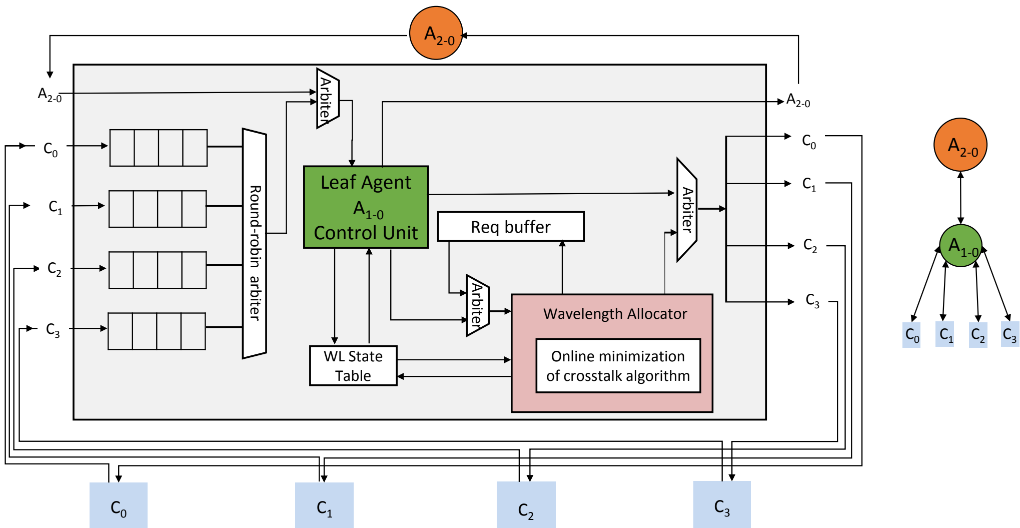


Figure 5.4: Illustration of detailed micro-architecture of leaf agent A_{1-0}.

- Request (line 3): it is first verified whether it reaches the first common ancestor agent of source and destination cluster. If NO, the request is forwarded to A_{2-0}. If YES, the request enters the wavelength allocator, where the selection of wavelengths is done with the help of online minimization of crosstalk algorithm detailed in section 5.2.3. Then an Ack is sent backward to the source cluster. If there is not enough free wavelengths, the request is forwarded to the Req buffer and will be retried after a certain delay.

- Ack (line 15): the source and destination ID are first compared to the local address in order to identify the related partition of wavelength table. For example, if an Ack reaches its ancestor leaf agent of source cluster, the partition from C_{src} to C_{N_C-1} in WL table should be freed, then the selected channel IDs in the rows between C_{src} and $C_{N_{in}-1}$ are updated from F to R. Finally the Ack is sent to the source cluster to turn on the specified lasers.

- Rel (line 22): it first releases the selected channel ID in the related partition of WL table. Then it is sent backward to the destination cluster to turn off the specified filters.

- WL_STATE_Req (line 29): firstly, it verifies whether the related partition of wavelength table is free and there exists enough free wavelengths. If YES, a

Algorithm 1 Control Unit of Leaf Agent

```
1: Function ControlUnitLeafAgent (Message Mess )
2:
3: if Mess.Type==Req then
4:    if Mess.D_ID == My_ID then
5:       // Message reaches the first common ancestor agent of source and destination cluster
6:       if Related partition of WL Table is free && EnoughFreeWL then
7:          MinimizationCrosstalkAlgorithm(Mess) //Select proper wavelengths
8:       else
9:          SendToWait_Buffer(Mess)
10:      end if
11:   else
12:      // I'm not the recipient, so send message to the upper agent
13:      SendtoUpperAgent(Mess)
14:   end if
15: else if Mess.Type==Ack then
16:   // Free related lines in WL Table and reserve selected WL ID
17:   UpdateWL_Table(Mess)
18:   if Mess.D_ID == MyA1_ID || Mess.S_ID == MyA1_ID then
19:      //Sent to source cluster for turning on laser and to destination cluster for turning on filters
20:      SendToCluster(Mess)
21:   end if
22: else if Mess.Type==Rel then
23:   // Release the selected WL ID in WL Table
24:   UpdateWL_Table(Mess)
25:   if Mess.D_ID == My_ID then
26:      //Sent to destination cluster for turning off filters
27:      SendToCluster(Mess)
28:   end if
29: else if Mess.Type==WL_STATE_Req then
30:   if Related partition of WL Table is free and EnoughFreeWL then
31:      // the related partition of WL Table becomes reserved
32:      UpdateWL_Table(Mess)
33:      Generate NewMess : WL_STATE_Rps with FreeWL
34:      SendtoUpperAgent(WL_STATE_Rps)
35:   else
36:      Generate NewMess : WL_STATE_Rps with NoFreeWL
37:      SendtoUpperAgent(WL_STATE_Rps)
38:   end if
39: else if Mess.Type==WL_STATE_Rel then
40:   // Release the related lines in WL Table
41:   UpdateWL_Table(Mess)
42: end if
```

WL_STATE_Rps with the common wavelengths state of related clusters is generated and sent to branch level agent. Meanwhile, the related partition of wavelength table is reserved so that other requests can not access this part of table. If NO, a WL_STATE_Rps with zero free wavelength is sent back to branch level agent.

- WL_STATE_Rel (line 39): it releases the related partition of WL table according to source and destination cluster ID.

5.2.2 Branch Agent

Figure 5.5 shows a part of example TbNoC with a detailed micro architecture of branch agent A_{2-0}. Similarly to leaf agent, 4 FIFOs are followed by a RR arbiter. Differently, the message should pass through a selector before reaching FIFO. If several WL_STATE_Rps enter branch agent at the same time, the AND operation for WL_STATE part is realized directly to find common WL_STATE and a new common WL_STATE_Rps is generated. Then the message reaches the control unit, the algorithm 2 is executed. Each type of message is managed differently as detailed below:

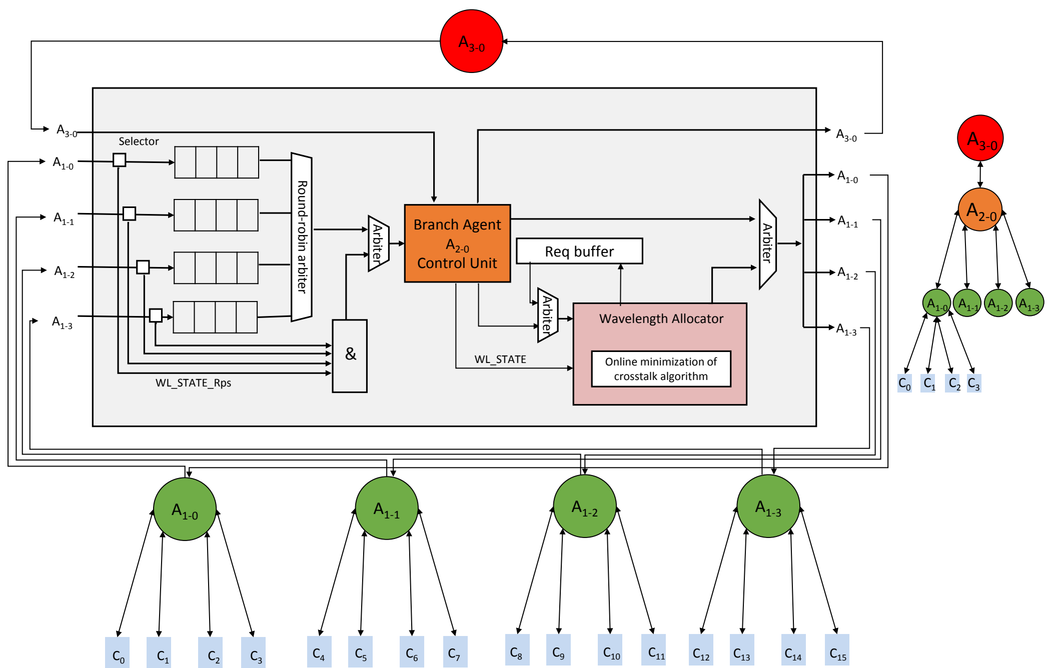


Figure 5.5: Illustration of detailed micro architecture of branch agent A_{2-0}.

- Request (line 3): it is first verified whether it reaches the common ancestor of source and destination cluster. If NO, the request is forwarded to the third level agent. If YES, the request is sent to wavelength allocator. However, the wavelengths allocation can not be launched directly due to the lack of wavelengths state informations. A WL_STATE_Req is thus generated and sent to the leaf agents.

Algorithm 2 Control Unit of Branch Agent

```
1:  Function ControlUnitBranchAgent (Message Mess )
2:
3:  if Mess.Type==Req then
4:    if Mess.D_ID == My_ID then
5:      // Message arrives at destination agent
6:      if EnoughFreeWL then
7:        WavelengthAllocation(Mess)
8:      else
9:        SendToWait_Buffer(Mess)
10:     end if
11:   else
12:     // I'm not the recipient, so send message to the upper agent
13:     SendtoUpperAgent(Mess)
14:   end if
15: else if Mess.Type==Rel || Ack || WL_STATE_Req || WL_STATE_Rel then
16:   //It is compared to local ID in order to identify the related section of lower agents
17:   SendtoRelatedLowerAgent(Mess)
18: else if Mess.Type==WL_STATE_Rps then
19:   //I'm not the recipient, so send message to the upper agent
20:   if Mess.D_ID! = My_ID || Mess.S_ID! = My_ID then
21:     SendtoUpperAgent(Mess)
22:   else
23:     if NoEnoughFreeWL then
24:       Generate NewMess : WL_STATE_Rel
25:       SendToRelatedLowerAgent(WL_STATE_Rel)
26:       SendToWait_Buffer(Req)
27:     else
28:       Generate NewMess : Ack
29:       SendToRelatedLowerAgent(Ack)
30:     end if
31:   end if
32: end if
33:
```

- Ack/Rel/WL_STATE_Req/WL_STATE_Rel (line 15): these messages should be forwarded to leaf agent and A_{2-0} is only a relay station. The source and destination ID are compared to the local address in order to identify the related leaf agents in the communication path.
- WL_STATE_Rps (line 18): if the branch agent is not the recipient, the packet is forwarded to the third level agent. Otherwise, this packet is forwarded to the wavelength allocator. The number of free wavelengths is first compared to the Req_NUM_WL in order to verify if there exists enough resource along the communication path. If No, a WL_STATE_Rel is generated to free WL table and cancel the allocation process; if Yes, a Ack is generated to validate the allocation.

The third level agent is similar to the second level agent, where WL table is not integrated. Differently, the third level agent is root agent which has no upper agent in our example.

5.2.3 Online Minimization of Crosstalk Algorithm

From MR transmission function detailed in chapter 2, it is clear that the choice of wavelengths for overlapping communications has a direct impact on the noise introduced at the photodetector. It is not hard to understand that the occurrence of a new wavelength might degrade the SNR performance of existing channel due to the additional crosstalk noise.

Algorithm 3 Online minimization of crosstalk algorithm

Function OnlineMinimizationCrosstalk (Req, $Common\ WL\_STATE$)
// Definitions :
// SelectedWL is the set of selected wavelength
// CalculateCost() is the objective function which computes $NN(\lambda_i)$
$SelectedWL = \{\}$
while $SelectedWL.length < Req\_Num\_WL$ **do**
  $Cost\_\lambda_{Best} = positive\ infinity$
  for $all\ \lambda_i\ in\ FreeWL$ **do**
    $\lambda_{Current} = \lambda_i$
    // The next function computes the normalised noise introduced by $\lambda_{Current}$
    $Cost\_\lambda_{Current} = CalculateCost(\lambda_{Current})$
    if $Cost\_\lambda_{Current} < Cost\_\lambda_{Best}$ **then**
      $\lambda_{Best} = \lambda_{Current}$
      $Cost\_\lambda_{Best} = Cost\_\lambda_{Current}$
    end if
  end for
  $SelectedWL.add(\lambda_{Best})$
  $FreeWL.remove(\lambda_{Best})$
  $ReservedWL.add(\lambda_{Best})$
end while

Our proposed algorithm comes from the same concept of its offline counterpart: allocating distinct and separate wavelengths for the optical signal. However, applying genetic algorithm for each request leads to high cost in time and area. In this context, we propose an online minimization of crosstalk algorithm. Firstly, we define the Normalized Noise (NN) introduced by a new wavelength λ_i as:

$$NN(\lambda_i) = \sum_{\forall \lambda_m \in ReservedWL} \frac{1}{|i-m|^2} \tag{5.3}$$

Where ReservedWL is the set of reserved wavelengths along the whole communication path, extracted from common WL_STATE. Then we follow the algorithm 3 to minimize the crosstalk noise. Our idea is to make the locally optimal choice at each step. Firstly, a wavelength that introduces the minimal crosstalk for existing channel is selected, then remove this selected wavelength from FreeWL and add it to ReservedWL, where FreeWL is the set of free wavelengths along the whole communication path. Finally repeat two steps above until enough wavelengths are selected. Consequently, a global suboptimal solution can be reached.

5.3 Hybrid opto-electronic NoC

TbNoC is designed to control optical channel set-up process for global communication in optical NoC. It is suitable for supporting different optical networks, such as Chameleon and SUOR, which demonstrates the high network flexibility.

5.3.1 Hybrid Network Combining Chameleon and TbNoC

Figure 5.6 illustrates a considered hybrid network composed of an optical layer implementing Chameleon and an electrical layer implementing a 2-level TbNoC with 8 clusters.

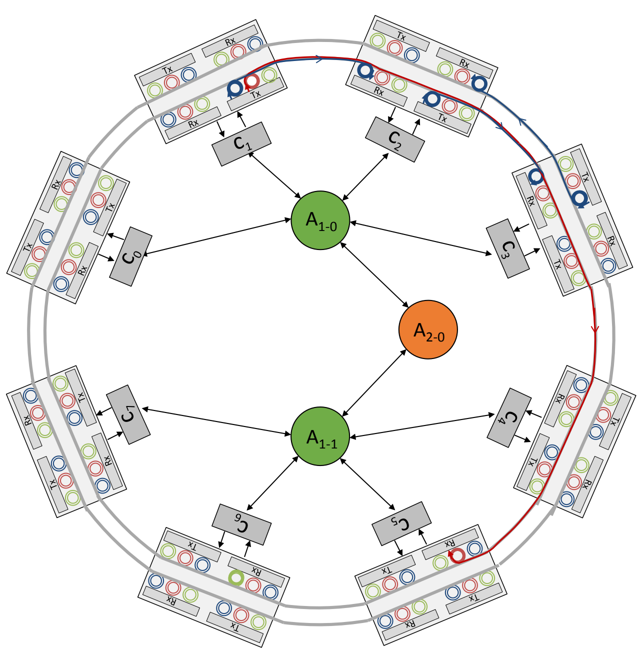


Figure 5.6: Overview of Chameleon and 2-level TbNoC hybrid NoC.

Each core cluster is connected to a given Optical Network Interface (ONI). The ONIs are crossed by two waveguides propagating optical signals. Each waveguide is single directional and two waveguides can be used to propagate optical signals. In Chameleon, dedicated point-to-point communication is facilitated by waveguide partitioning, which allows to realize multiple independent communications on a given wavelength. In figure 5.6, blue signal is used to realize communication $C_1 \rightarrow C_2$ and $C_2 \rightarrow C_3$ at the same time, these two optical path set-ups are realized in A_{1-0}. Thanks to WDM, each wavelength corresponds to a physical channel. Hence, red signal for $C_1 \rightarrow C_5$ has no impact on blue signal. This long-distance channel set-up process is handled by A_{2-0}. Moreover, the combined use of clockwise and counter-clockwise directions for signal propagation allows a substantial improvement of its energy efficiency. In this example, blue signal in another waveguide is used for $C_3 \rightarrow C_2$ in counter-clockwise direction.

5.3.2 Hybrid Network Combining SUOR and TbNoC

Figure 5.7 illustrates a considered hybrid network composed of an optical layer implementing SUOR and an electrical layer implementing a 2-level TbNoC with 8 clusters.

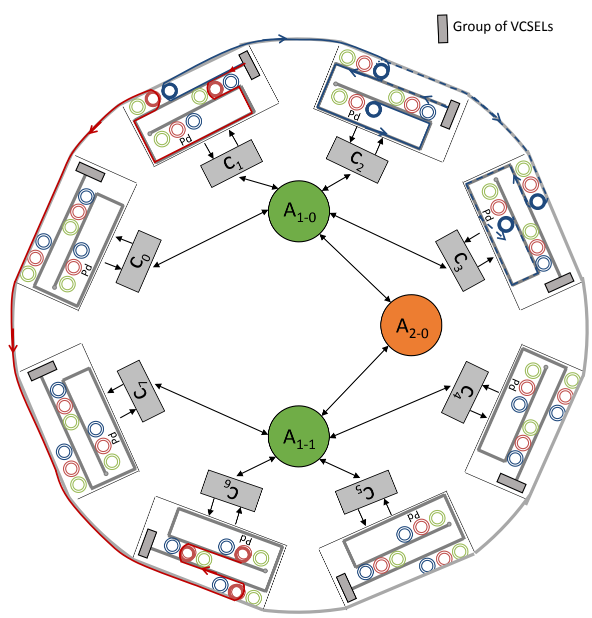


Figure 5.7: Overview of SUOR and 2-level TbNoC hybrid NoC.

In SUOR, one waveguide can support bidirectional transmissions. For example, blue signal is used for communication $C_1 \rightarrow C_2$ in clock direction while red signal is used for communication $C_1 \rightarrow C_6$ in counter clock direction. Similarly, SUOR also supports waveguide partitioning, which allows blue signal (for the sake of clarity, in dotted line) to realize another communication $C_2 \rightarrow C_3$ at the same time.

5.4 Results

In this section, we first present 28GH FDSOI synthesis results of TbNoC. Then we explore the design space in terms of the number of tree level and the number of wavelengths per waveguide. Finally, we compare our allocation mechanism with other three mechanisms introduced in SUOR, Corona, and Chameleon-Baseline [58].

5.4.1 28nm FDSOI Synthesis Results

We implemented a VHDL model of leaf agent and the synthesis results are shown in table 5.3. Firstly, synthesis of the behavioural-level of leaf agent into a gate-level verilog netlist is performed on 28nm FDSOI technology by using Synopsys Design Vision environment (Version H-2013.03-SP5). The 28nm FDSOI CMOS libraries provide definitions, logical descriptions and timing information of various logical gates. Then in order to verify the correctness of produced hardware during the synthesis process and its correspondence to the initial HDL description, we create a test bench by using the input data of a target leaf agent extracted from the simulation of neighbor traffic running on 64-clusters TbNoC, with a injection rate 10 Reqs/clk. Additionally, we use ModelSim to produce Value Charge Dump (VCD) file during simulation process to log switching activity needed for power estimation [86]. Finally, integrated gate-level power estimation is performed by means of Synopsys PrimeTime PX suite that requires the netlist data, switching activity recorded by the VCD file and 28nm FDSOI library specifications as input parameters. Synthesis results show, area, critical path and power consumption according to the number of wavelengths in one waveguide. Vdd is set to 1V and the working frequency is set to 100MHz. In general, the areas and power consumption go up with the number of wavelengths. It can be seen that one leaf agent requires 106364μm^2 for 64 wavelengths, while consuming 9.34 mW for the aforementioned configuration. 74.6 % reduction in energy overhead is obtained when 8 wavelengths gather in the waveguide. Dynamic power consists of internal power and switching power which is predominant in power consumption. The synthesis results clearly show that our architecture can be efficiently implemented in 28nm FDSOI technology.

Table 5.3: Leaf agent synthesis results.

Number of Wavelength	Area μm^2	Critical Path (ns)	Power Consumption			
			Switching Power (mW)	Internal Power(mW)	Leakage Power (mW)	Total (mW)
8	16922	0.46	0.07	2.07	0.23	2.37
16	28691	0.81	0.11	2.77	0.41	3.29
32	62220	1.33	0.4	4.24	0.9	5.54
64	106364	2.43	0.8	7.11	1.43	9.34

5.4.2 Simulation Results

The simulation is executed in VHDL and performed under different synthetic traffics. Before introducing the performance metrics for system, it is useful to consider the spatial distribution of message in interconnection networks. Table 5.4 lists some static traffic patterns used in this section.

Traffic names	Details
Uniform	Each cluster sends packets randomly to other clusters
Hot Spot	All clusters send packets to one specific cluster
Neighbor	Each cluster sends packet to the 1 hop neighbor

Table 5.4: Synthetic traffic patterns.

We evaluate the network based on its set-up latency. It includes the waiting time for a request in the queue and the set-up delay for a request, which extends from first time sending until the final success. Data size refers to the duration of data transmission multiplied by the data rate per wavelength and the requested number of wavelengths. Therefore we can use the duration of optical path to represent the data size when we assume the data rate per wavelength and the requested number of wavelengths for each communication are the same. The duration of data transmission and injection rate of set-up requests are two main factors for the set-up latency.

5.4.2.1 Architectural Exploration

Number of tree levels: in this section, we investigate impact of the number of tree levels on average set-up delay. 3-level tree for 64-clusters network has been detailed in the section 5.1. 2-level and 4-level trees for 64-clusters network are illustrated in figure 5.8 (a) and figure 5.8(b) respectively.

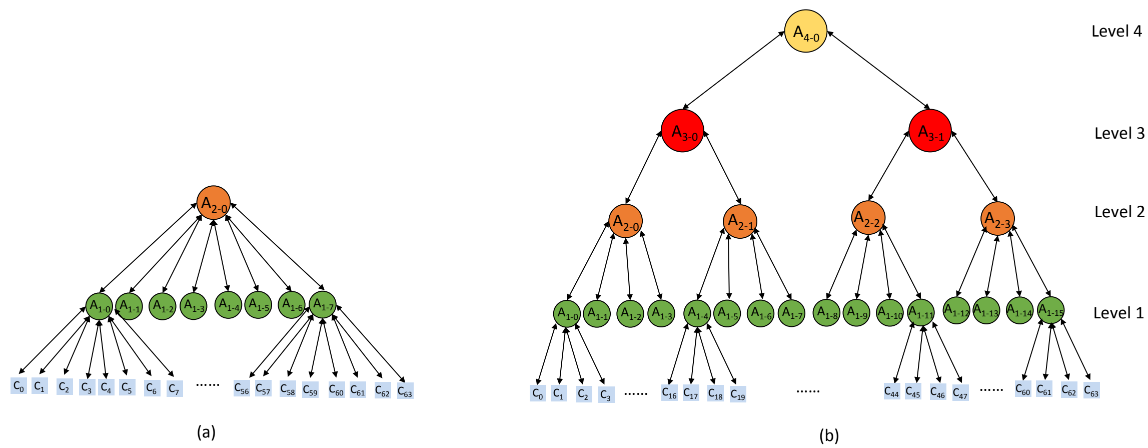


Figure 5.8: (a) 2-level (b) 4-level tree-based control network.

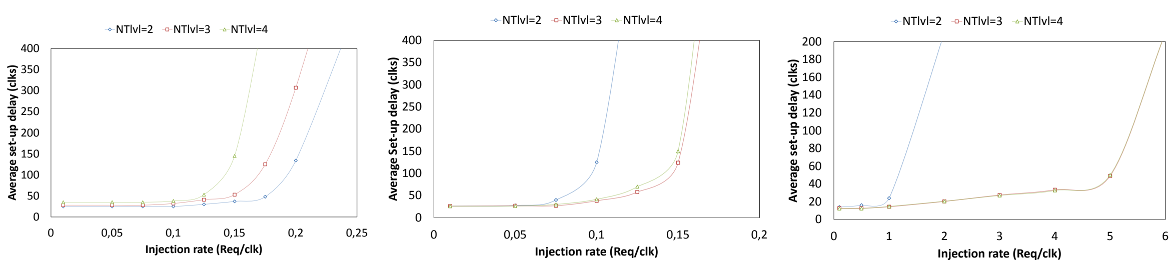
We assume a waveguide and 16 wavelengths per waveguide. The duration of data is set to 50 cycles and 1 wavelength is requested for each communication. Figure 5.9 (a) shows the set-up latency under uniform traffic pattern. Here, 1 req/clk means 1 req per cycle in the network. The average set-up delay increases with injection rate, since the delay incurred by contention is proportional to injection rate. Thus, the greater the injection rate, the higher penalty of contention. The network can only work stably with injection rate smaller than the saturation point. The delay climbs dramatically after saturation point for all designs. Besides, 2-level tree outperforms other two architectures. For example, the saturation point in 2-level tree is 0.24 compared to 0.2 and 0.17 for 3-level tree and 4-level tree respectively.

Figure 5.9(b) shows the set-up delay under hotspot traffic pattern, in which case all clusters transmit packets to a specific cluster. The concentration of traffic sharply increases the set-up delay of each architecture. Generally speaking, the 2-level tree goes up more quickly than other two architectures, since the root agent in 2-level tree is very busy due to the concentration of traffic.

Figure 5.9(c) shows the set-up delay under neighbor traffic pattern, simulating the well-mapped tasks with communications with high locality. Totally, the saturation point in neighbor traffic is higher than two others traffics. 1.5 req/clk compares to 0.24

req/clk and 0.12 req/clk for uniform and hotspot traffic respectively when we simulate 2-level tree. The performance gain under neighbor traffic shows that TbNoC supports locality quite well. This is due to the channel partition and high resource flexibility. For instance, each channel can be divided into 64 sections and each section can serve as a channel for the 1 hop neighbor transmission. However, 2-level tree performs worse, since only 8 leaf agents are implemented in the control network, thus fewer local transmission requests can be handled in parallel.

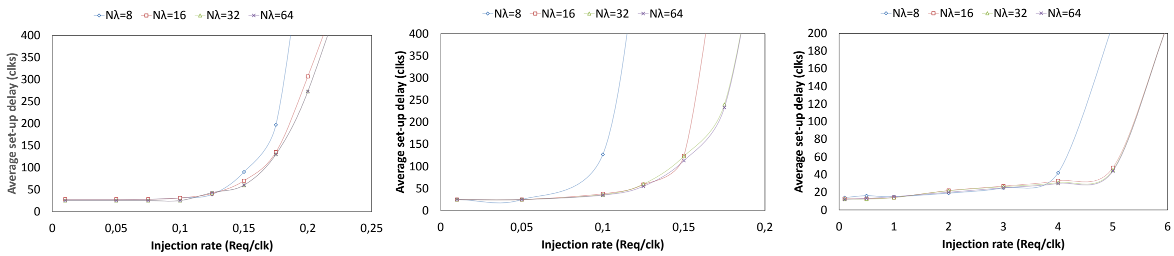
In summarize, the smaller the number of tree levels, the higher the number of input ports and the more fierce competition of limited link resources. Nevertheless, a smaller number of network level can reduce the path set-up latency for long-distance communication request. Therefore, the proper number of levels is important for the our design of efficient network.



(a) (b) (c)

Figure 5.9: The average set-up delay for various number of tree level under different traffic patterns (a) Uniform (b) Hotspot (c) Neighbor.

Number of Wavelengths: in this study, we investigate impact of the number of wavelengths on the average set-up latency. For this purpose, the simulation is performed on a 3-level tree and each communication requests one wavelength. Figure 5.10 shows the simulation results for the number of wavelengths 8, 16, 32 and 64 under different traffic patterns. Generally, the saturation point rises as the increase of the number of wavelengths. For example, the saturation point is improved by 1.5 times between $N_\lambda = 8$ and $N_\lambda = 16$ under hotspot traffic. This is due to more available wavelengths to alleviate the conflict of optical resource. However, the performance gains cannot be continuously improved as the number of wavelengths increases. When 32 and 64 wavelengths are employed, the saturation points are nearly same. Since the limited link resources become the predominant factor in set-up delay.



(a) (b) (c)

Figure 5.10: The average set-up delay for various number of wavelengths under different traffic patterns (a) Uniform (b) Hotspot (c) Neighbor.

5.4.2.2 Performance Evaluation

In this section we compare TbNoC with other wavelength allocation mechanisms developed in different architectures: Corona, SUOR, and Chameleon-Baseline [58] in terms of SNR and set-up delay under different synthetic traffics. We assume the same optical interface, Chameleon for these four control networks. To calculate the SNR of application running on optical layer, we take use of the same power model presented in chapter 2. The architectural parameters are given in table 5.5.

Table 5.5: Technological parameters.

Parameters	Value
Propagation loss	-0.274dB/cm
Photodetector sensibility	-20dB
Laser optical output power	0dBm
Resonance shift	0.4nm
FSR	96nm
Number of wavelengths per waveguide	64
$-3dB$ bandwidth of MR	0.26nm

Figure 5.11 illustrates wavelengths arbitration protocols in TbNoC and the other three alternative designs. In SUOR, the cluster sends request to its agent to set up optical channel. To cope with the complexity of arbitration in cluster agent, channel grouping is used to avoid conflict between inter-group transactions. The data channels

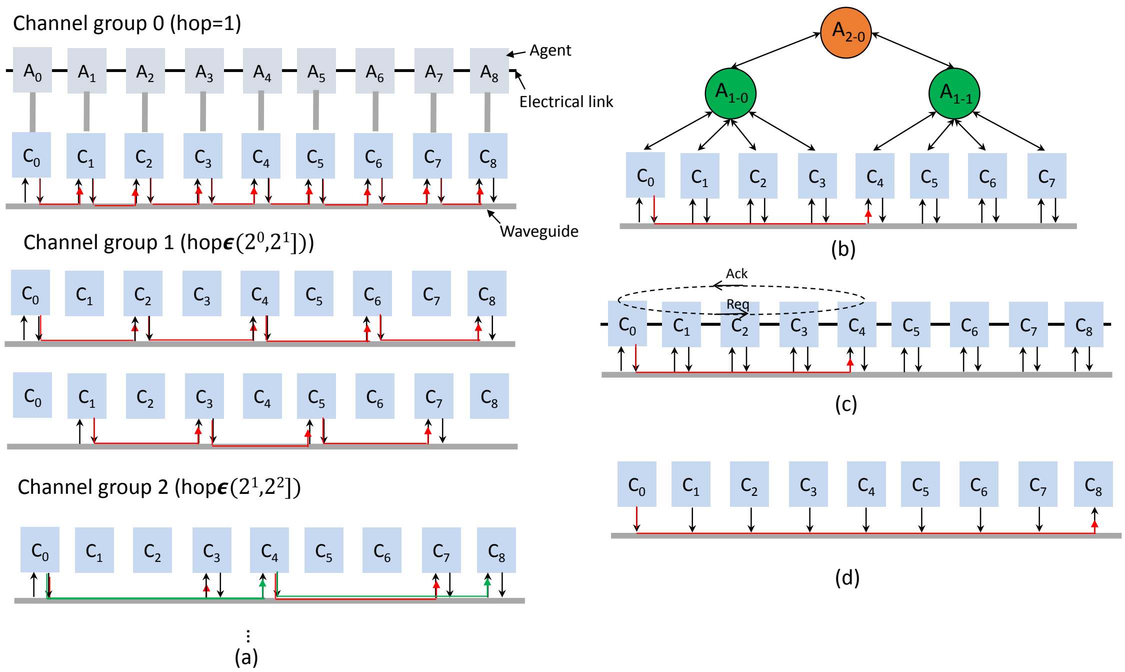
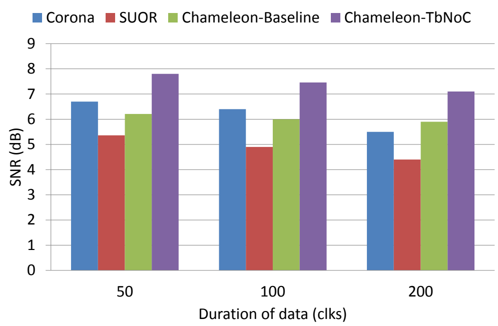
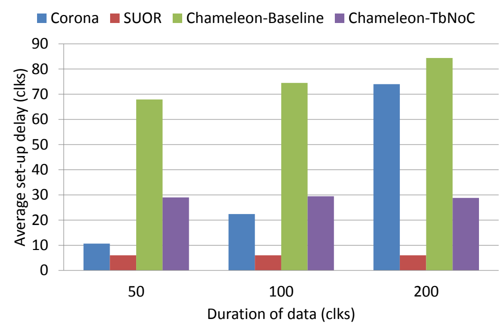


Figure 5.11: Illustration of wavelength arbitration methods of (a) SUOR (b) Chameleon-TbNoC (c) Chameleon-Baseline (d) Corona. The colored lines with arrows show the current transmission on the channel. The black arrow show the access licence of channel as writer or reader.

are classified into different groups according to the number of hops between source and destination. In each group channel, only the transactions sharing a same group of distance are allowed. As shown in figure 5.11 (a), in the i-th group, only transactions with distances $\in (2^{i-1}, 2^i]$ are allowed. In TbNOC, a tree-based control network is use to handle the requests of accessing the data channel. Each agent employs a centralized control mechanism and governs a certain section of clusters. In Chameleon-Baseline, the control issue is handled by a packet-switched electrical network as shown in figure 5.11 (c). The control packet may pass through multiple routers without buffered as long as no collision happens. Each request tries to communicate firstly on first free wavelength. If it failed, a Nack will be sent back and a new request will try on the next free one, and so on. Corona employs MWSR crossbar, where each cluster reads from a dedicated data channel while all clusters can write to this channel as shown in figure 5.11 (d). Generally, imposing access rules may sacrifice some flexibility of the network, in turn, it effectively reduces the arbitration overhead. For fair comparison, we assume one waveguide gathering 64 wavelengths in the optical layer. Moreover, each data channel is composed by only one wavelength, instead of 4 waveguides as assumed in Corona [9] and one waveguide in SUOR [10]. For SUOR, we keep the same channel grouping method but reduce the number of data channels in each group due to the limited available resources (here 64 wavelengths instead of 284 waveguides as assumed in SUOR).



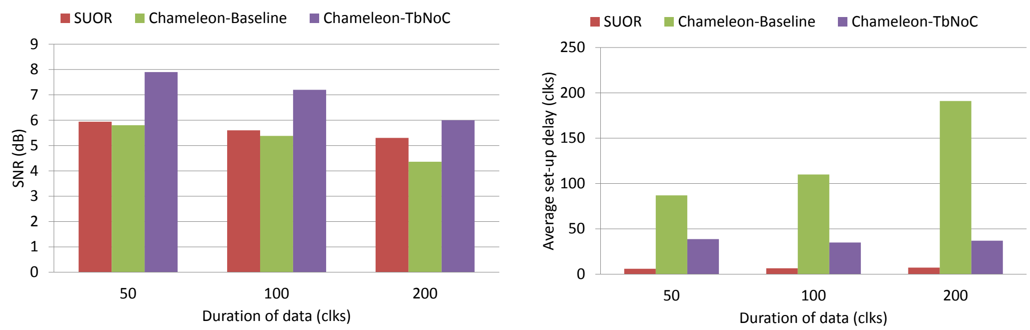
(a)



(b)

Figure 5.12: The comparisons of (a) SNR (b) average set-up delay for 4 mechanisms under uniform traffic with different duration of data.

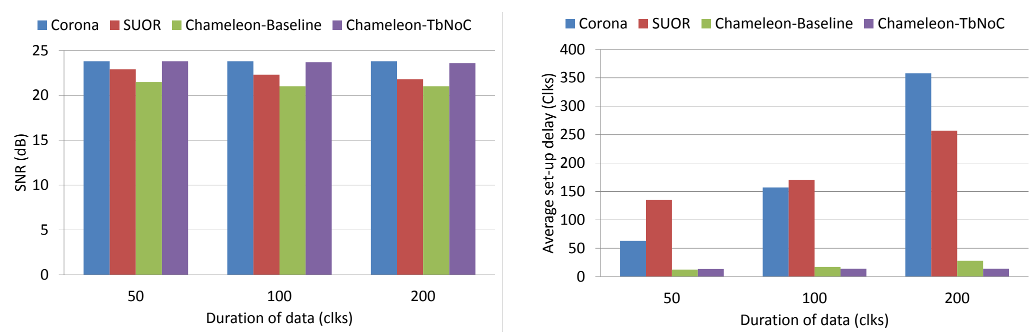
As shown in figure 5.12 (a), TbNoC has a better SNR compared with others three mechanisms under uniform traffic with injection rate equals to 0.1. When duration of data equals to 50clks, TbNoC obtains a SNR of 7.8dB, compared to 6.8dB for Corona, 5.2dB for SUOR, and 6.2dB for Chameleon-Baseline. This is due to the high channel flexibility and the minimization of crosstalk algorithm implemented in TbNoC. When the data duration increases, the SNR decreases due to increasing parallel optical signals in the waveguide. Figure 5.12 (b) shows the comparison of average set-up delay. In Corona, the set-up delay increases with the data duration. Because only one transaction is allowed at a time on a single data channel. For Chameleon-Baseline, the penalty of contention is high, because it needs to tear down all the reserved resources along the path and then the source re-emits the reservation message after a random latency. Hence, Chameleon-Baseline demonstrates a bad set-up delay performance. The hierarchical central control mechanism in TbNoC and channel grouping rule in SUOR work well under different data duration in this low loaded uniform traffic, 29clks for TbNoC and 8clks for SUOR. SUOR demonstrates a smaller set-up delay due to low-complex arbitration strategy, it sacrifices some flexibility of network and effectively reduces the arbitration overhead.



(a) (b)

Figure 5.13: The comparisons of (a) SNR (b) average set-up delay for 4 mechanisms under hotspot traffic with different duration of data.

Figure 5.13 shows the network performance under hotspot traffic pattern with different duration. The injection rate is 0.1 request per clock cycle. The results of Corona are not illustrated in this figure due to its high set-up latency (over 2000 clks). All the requests require only one specific wavelength under hotspot traffic according to its access rule. Thus Corona is not recommended for hotspot traffic. Besides, the concentration of traffic degrades also greatly the latency performance of Chameleon-Baseline.



(a) (b)

Figure 5.14: The comparisons of (a) SNR (b) average set-up delay for 4 mechanisms under neighbor traffic with different duration of data.

Figure 5.14 shows the network performance under neighbor traffic pattern with different duration. The injection rate is 1 request per clock cycle. The SNR performance in neighbor traffic is better compared to two traffics above. For example, 23.8dB under neighbor traffic compared to 8dB under uniform and hotspot traffic for TbNoC. Since 1-hop communications suffer from less insertion loss along path and less inter-channel crosstalk from others overlapping communications. This demonstrates that the importance of placing high-communicating tasks on the neighbor cores. However, the average set-up delay for Corona and SUOR are very high. In fact, Chameleon-Baseline and TbNoC work well in the high load neighbor traffic due to the high channel flexibility. Corona and SUOR are limited by the access rules in the data channel, this disadvantage is exposed in high load traffic. For Corona, each cluster can read from one specific data channel. For SUOR, the communications within a certain distance share the same group of data channels. Thus only a small portion of data channels are used for busy 1-hop communication. When a high load neighbor traffic runs on Corona or SUOR, a request often needs to wait in source cluster until the previous request finishes transmission.

5.5 Conclusions

Online wavelength allocation is necessary to support dynamic management of optical resource. In this chapter, we propose an electrical tree-based network, named TbNoC to handle online optical channel set-up requests. TbNoC is generic and can be combined to different ONoC based on shared waveguide, such as Chameleon and SUOR. Multiple characteristics are designed to improve power efficiency and system performance. First of all, the global electrical control network is divided into several hierarchical central-controlled agents, thus optical channel set-up delay may be reduced. Secondly, a minimization of crosstalk algorithm, which allocates distinct and separate wavelengths for optical signals, is employed to boost SNR performance as well as to reduce power consumption. Finally, each channel can be accessed by all the clusters which improves the flexibility of network. As a result, we explore the design space of TbNoC in terms of the number of tree levels and the number of wavelengths. To illustrate the SNR of TbNoC, we instantiate TbNoC to Chameleon ONoC, named Chameleon-TbNoC. Then we show that Chameleon-TbNoC improves a SNR of 1.6dB compared to SUOR under uniform traffic. Since the fully shared resource mechanism in TbNoC is more flexible compared to the limited access rules implemented in SUOR, but it results in a higher complexity of network. Consequently, a design trade-off is that opening channels flexibility leads to heavy arbitration burden and high network complexity, while lacking flexibility would cause the resources waste and SNR performance degradation.

5.6 Comparison of Offline and Online Approaches

This thesis focus on two ways to allocate wavelength in ONoC: offline and online. In the former case, the communication requests are known in advance, then an extra offline simulation is required to select proper wavelengths. While in the latter case, a sequence of communication requests is received at run-time and an on-demand optical path is set-up in response to each request.

Table 5.6: Comparison of online and offline approaches for wavelength allocation problem in photonic interconnect.

Comparison	Offline	Online
Application known in advance	Yes	No
The optimum results	Yes	No
Communication protocol complexity	High	Low
Set-up latency	Zero	Limited by ongoing communications
An extra offline simulation	Yes	No
Static mapping	Yes	No

In general, online approach cannot match the performance of offline approach due to the incomplete knowledge of the whole application input. It is clear that the occurrence of a new wavelength might degrade SNR of existing channel due to the added crosstalk noise. Thus the lack of future communication is a great drawback for online approach. Moreover, the process of set-up path introduces an extra latency for each communication compared to their offline counterpart. However, offline approach needs an extra offline simulation before running application on ONoC. Furthermore, offline approach should know the mapping of tasks in advance for offline approach, while online approach allows the tasks to remap onto the cores according to specific requirement during the execution.

As we know, one main concern of optical network is the limited on-chip power budget for current MPSoC. The higher the number of wavelengths per channel, the higher the crosstalk and the higher the laser output power needed to satisfy specific

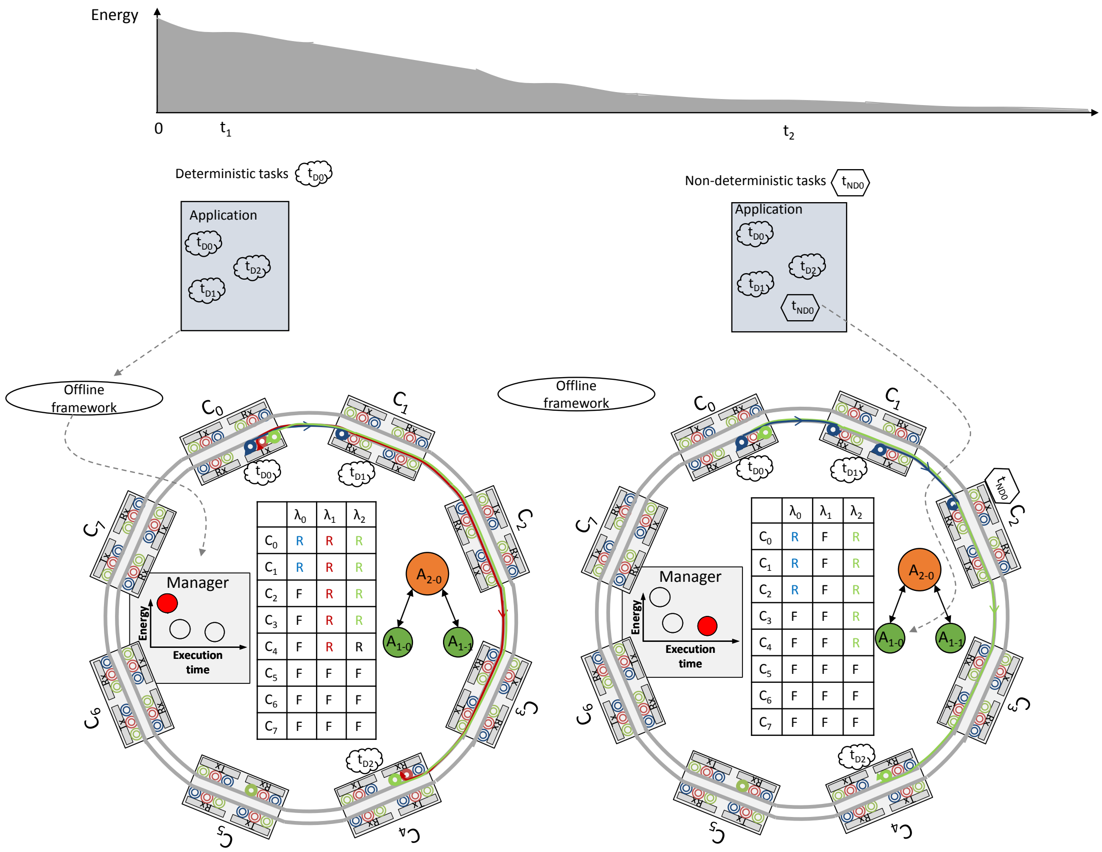


Figure 5.15: Unified offline and online wavelength allocation approaches.

BER requirement. This leads to the following conflicting objectives: high performance communications tend to rely on an exhaustive use of the available wavelengths while energy efficient communications involve a parsimonious use of signals. Hence, a proper number of wavelengths should be allocated according to execution requirement (e.g. high performance and low power). For instance, a constraint may be an execution deadline for real time systems, high performance solution for which multiple wavelengths are allocated to shorten the communication time is considered. Then the battery of system is nearly ran out, low power solution which tend to minimize the number of used wavelengths is selected. Moreover, when mixed deterministic and non-deterministic tasks occur, offline and online wavelength approaches should be jointly considered to ensure the efficient communication.

As shown in figure 5.15, several deterministic tasks appear at t_1, their wavelengths are solved by our proposed offline framework at compile time, which results in a set of offline configurations stored in a specific memory. Then configuration manager can deploy it at run-time according to execution requirement. In the beginning, the abundant energy supply can drive the application runs as fast as possible. Hence, the manager takes use of high performance configuration mode (λ_0, λ_1, and λ_2 are reserved). At t_2, a new non-deterministic task t_{ND0} appears, leading to more data exchange in the network. Moreover, a lot of energy are consumed during previous application executions. Hence, the energy-saving mode should be considered and the manager change to load low power configuration mode for the deterministic tasks. The wavelength table is updated according to selected configuration mode (λ_0 and λ_2 are reserved). At the same time, the non-predictive task t_{ND0} is forwarded into TbNoC (here we assume A_{1-0} can handle this task), where online channel set-up process is launched to allocate wavelengths from those which are not used by offline configuration. In summarize, our offline and online approach can be combined to resolve wavelength allocation problem in some realistic application, such as aircraft detection.

Chapter 6

Conclusions and Perspectives

6.1 Overview

Increasing high bandwidth on-chip communication demands among cores put a lot of pressure on traditional electric NoCs. With the recent progress in silicon photonic, Optical NoC stands out as one of most promising emerging technology for on-chip communication due to its high bandwidth and low data-dependant power dissipation. A typical silicon photonic link is comprised of: 1) a laser source to emit optical signal 2) a waveguide to propagate optical signal, and 3) a photodetector to recover electrical signal. WDM are employed in silicon waveguide to open high bandwidth channel for data transmission. However, optical signals of various wavelengths can interfere with each other through optical switching elements, introducing crosstalk noise. The higher the number of signals propagating simultaneously, the higher the crosstalk and the higher the laser output power required. This leads to the contradictory objectives related to BER requirements, expected execution performance and power budget.

In this thesis, we assess the performance and energy efficiency of ONoC from offline and online wavelength allocation respectively. In the former case, the communication needs are known in advance; while in the latter case, the optical paths are set-up on demand at run-time.

Firstly, we proposed an offline methodology based on genetic algorithm allowing the exploration of performances and energy consumption trade-off for ONoC-based MPSoC. In order to find the best trade-off, wavelength allocation and laser power scaling are combined to configure system. The methodology is formed as a framework handling several inputs, such as application, architecture, and device parameters. On the one hand, the exploration of design space at both device level and system level is allowed; on the other hand, a set of performance and energy efficiency trade-offs can be generated once the input parameters are defined. The resulting Pareto solutions include 1) low-power solutions, which tend to minimize the number of used wavelengths, and 2) high-performance solutions, for which multiple wavelengths are allocated to shorten the communication time. As an example, for a 63-task application running on 64 cores, the relative variation in the execution time and energy is 60% and 39% respectively. For a 107-task application running on 256 cores, the relative variation in the execution time and energy is 69% and 40% respectively. Finally, these solutions can be embedded in a controller and deployed at run-time according to execution requirements (e.g. high performance, low power and application constraints like deadlines).

Then, this work also explores the online wavelength allocation. We proposed an electrical tree-based NoC, named TbNoC, to handle online optical channel set-up requests. TbNoC is generic and can be combined to different ONoCs based on shared waveguide, such as Chameleon and SUOR. In this control network, several characteristics are designed to improve power efficiency and system performance. First of all, the global electrical control network is divided into several hierarchical central-controlled agents, thus optical channel set-up delay may be reduced. Secondly, a minimization of crosstalk algorithm, which allocates distinct and separate wavelengths for optical signals, is employed to boost SNR performance as well as to reduce power consumption. Finally, each channel can be accessed by all the clusters which improves the flexibility of network. All of these techniques help to build the high-performance power-efficient control network. As a result, we explore the design space of TbNoC in terms of the number of tree levels and the number of wavelengths. To illustrate the SNR of TbNoC, we instantiate TbNoC to Chameleon ONoC, named Chameleon-TbNoC. Then, we show that Chameleon-TbNoC improve a SNR of 1.6dB compared to SUOR under uniform traffic. While the fully shared resource mechanism in TbNoC is more flexible compared to the limited access rules implemented in SUOR, it results in a higher complexity of network. Consequently, a design trade-off is that opening channels flexibility leads to heavy arbitration burden and high network complexity, while lacking flexibility would cause the resources waste and SNR performance degradation.

6.2 Perspectives

6.2.1 Online Laser Power Scaling

To minimize overall power consumption, there is a clear need for energy laser management. This means that with lower channel utilization, the laser power needs decrease, and the global power consumption decreases proportionally. In this case, a dynamic fine-grained control strategy for minimizing laser power consumption while meeting the communication bandwidth requirement is needed. We propose to configure the laser current to generate multiple levels laser power such that a central manager can intelligently switch among different levels based on ongoing communications. A similar technique is included in our offline approach, but for online management an heuristic must be defined to evaluated the trade-off between laser power consumption and communication BER.

6.2.2 Task Mapping and Scheduling for ONoC

With hierarchical network design, communication systems have adapted by putting an extra emphasis on creating locality and low-cost local electrical network in ONoC which is vital for performance improvement and power reduction. For example, in hot-spot traffic, one can consider that highly-communicating tasks should be placed in a same cluster. In this case, the most hot-spots can be managed at cluster level and not at ONoC level. Then optical network providing high bandwidth is used for long-distance inter-cluster communications. As we know it, the crosstalk introduced by conflicting wavelengths from different communications can be avoided by moving tasks in space and time. Besides, mapping a task on a core which is nearest to the destination core can introduce lowest losses along the optical path. We propose to map the tasks dynamically on the cores such that energy consumption is reduced. When there is a new task to be mapped onto the cores, it is first passed to the arbitrated manager. This arbitrated manager has a table known as mapping table, which is shared by all the IP cores. It will check in the mapping table which cores are free to execute the task. If there exists a free core located on the same cluster as the destination core does, this neighboring core is selected. Otherwise, for each free core, the number of hops to the destination core and the number of used wavelengths for ongoing communications along the path are calculated. Finally, the task is mapped on a core which leads to the smallest losses and crosstalk.

6.2.3 Thermal and Process Variation

Two key technical challenges facing photonic interconnect are the sensitivity of photonic devices to process and temperature variations. The existence of inherent spatial variations in CMOS materials and fabrication processes degrades more greatly silicon photonic devices than electronic devices [33]. In our BER estimation model, we calculate the losses experienced by the all the transmitting components, from laser sources to destinations, taking into account the ON-OFF states of the crossed MRs, but the dynamic thermal and static process variation effects of MR are not considered. However, the temperature sensitivity of nanophotonic devices and manufacturing process variation can cause resonant wavelength shift. This shift increases communication errors and bandwidth losses.

Every degree of temperature gradient between a ring modulator and ring filter leads to a 0.078nm mismatch in resonant wavelength [87]. Hence, we propose to model the laser source as a function of temperature and update laser power for each configuration at run-time through the data volume in the waveguide.

The process variation depends on several factors, including waveguide geometry, material and the fabrication process. A 1nm deviation in waveguide width leads to 0.76nm resonant wavelength shift [88]. When a resonance shift is over 1/3 of the channel spacing, an increase from 10^{-12} to 10^{-6} in BER can be introduced [89]. In this case, a mechanism is needed to estimate resonance wavelength shifts induced by process variation. For example, VARIUS tool is adapted to model die-to-tdie as well as within-die process variation in MRs [90]. Then we can allocate wavelength based on accurate resonant wavelengths of all the MRs across a chip to reduce the effect of process variation.

Bibliography

[1] B. K. Ajay Joshi and V. Stojanovic, “Designing energy-efficient low diameter on-chip networks with equalized interconnects.” Presented at 2009 17th IEEE Symposium on High-Performance Interconnects, 2009.

[2] L. P. Carloni, P. Pande, and Y. Xie, “Networks-on-chip in emerging interconnect paradigms: Advantages and challenges,” in *2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*, May 2009, pp. 93–102.

[3] M. Nikdast, “Signal-to-noise ratio in optical interconnection networks: Analysis, modeling, and comparison,” Ph.D. dissertation, The Hong Kong University of Science and Technology, 2013.

[4] C. Sciancalepore, B. B. Bakir, X. Letartre, J. Harduin, N. Olivier, C. Seassal, J. M. Fedeli, and P. Viktorovitch, “Cmos-compatible ultra-compact 1.55-m emitting vcsels using double photonic crystal mirrors,” *IEEE Photonics Technology Letters*, no. 6, pp. 455–457, March 2012.

[5] W. Bogaerts, L. Liu, e. I. Roelkens, Gunther", and G. Nicolescu, *Technologies and Building Blocks for On-Chip Optical Interconnects.* New York, NY: Springer New York, 2013, pp. 27–78. [Online]. Available: https://doi.org/10.1007/978-1-4419-6193-8_2

[6] L. Vivien, D. Marris-Morini, J. Mangeney, P. Crozat, E. Cassan, S. Laval, J. M. Fedeli, J. F. Damlencourt, and Y. Lecunff, “42 ghz waveguide germanium-on-silicon vertical pin photodetector,” in *2008 5th IEEE International Conference on Group IV Photonics*, Sept 2008, pp. 185–187.

[7] M. Briere, B. Girodias, Y. Bouchebaba, G. Nicolescu, F. Mieyeville, F. Gaffiot, and I. O’Connor, “System level assessment of an optical noc in an mpsoc platform,” in *2007 Design, Automation Test in Europe Conference Exhibition*, April 2007, pp. 1–6.

[8] P. Grani and S. Bartolini, "Simultaneous optical path-setup for reconfigurable photonic networks in tiled cmps," in *2014 IEEE Intl Conf on High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC,CSS,ICESS)*, Aug 2014, pp. 482–485.

[9] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. G. Beausoleil, and J. H. Ahn, "Corona: System implications of emerging nanophotonic technology," in *ACM SIGARCH Computer Architecture News*, vol. 36, no. 3. IEEE Computer Society, 2008, pp. 153–164.

[10] X. Wu, J. Xu, Y. Ye, Z. Wang, M. Nikdast, and X. Wang, "Suor: Sectioned undirectional optical ring for chip multiprocessor," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 10, no. 4, p. 29, 2014.

[11] kalray, "Mppa manycor," 2012.

[12] M. J. R. Heck and J. E. Bowers, "Energy efficient and energy proportional optical interconnects for multi-core processors: Driving the need for on-chip sources," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 4, pp. 332–343, July 2014.

[13] U. Y. Ogras and R. Marculescu, *Modeling, Analysis and Optimization of Network-on-Chip Communication Architectures.* Springer Publishing Company, Incorporated, 2015.

[14] S. LIU, "New circuit switching techniques in on-chip networks," Ph.D. dissertation, KTH School of Information and Communication Technology, 2015.

[15] P. T. Wolkotte, G. J. M. Smit, N. Kavaldjiev, J. E. Becker, and J. Becker, "Energy model of networks-on-chip and a bus," in *2005 International Symposium on System-on-Chip*, Nov 2005, pp. 82–85.

[16] A. B. Abdallah, *Multicore Systems On-Chip: Practical Software/Hardware Design.* Atlantis Press, Paris, 2013.

[17] N. Magen, A. Kolodny, U. Weiser, and N. Shamir, "Interconnect-power dissipation in a microprocessor," in *Proceedings of the 2004 international workshop on System level interconnect prediction.* ACM, 2004, pp. 7–13.

[18] A. Karkar, T. Mak, K. F. Tong, and A. Yakovlev, "A survey of emerging interconnects for on-chip efficient multicast and broadcast in many-cores," *IEEE Circuits and Systems Magazine*, vol. 16, no. 1, pp. 58–72, Firstquarter 2016.

[19] K. Salah, "Tsv-based 3d integration fabrication technologies: An overview," in *2014 9th International Design and Test Symposium (IDT)*, Dec 2014, pp. 253–256.

[20] Y. Qiu, S. Yue, W. A. Moussa, and P. Mousavi, "Development of embedded redistribution layer-based silicon interposer for 3-d integration," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 8, no. 3, pp. 399–409, March 2018.

[21] J. Kim, C. Nicopoulos, D. Park, R. Das, Y. Xie, V. Narayanan, M. S. Yousif, and C. R. Das, "A novel dimensionally-decomposed router for on-chip communication in 3d architectures," *SIGARCH Comput. Archit. News*, vol. 35, no. 2, pp. 138–149, June 2007. [Online]. Available: http://doi.acm.org/10.1145/1273440.1250680

[22] Y. Yano, T. Sugiyama, S. Ishihara, Y. Fukui, H. Juso, K. Miyata, Y. Sota, and K. Fujita, "Three-dimensional very thin stacked packaging technology for sip," in *52nd Electronic Components and Technology Conference 2002. (Cat. No.02CH37345)*, 2002, pp. 1329–1334.

[23] V. F. Pavlidis and E. G. Friedman, "3-d topologies for networks-on-chip," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 10, pp. 1081–1090, Oct 2007.

[24] D. A. B. Miller, "Rationale and challenges for optical interconnects to electronic chips," *Proceedings of the IEEE*, vol. 88, no. 6, pp. 728–749, June 2000.

[25] D. DiTomaso, A. Kodi, D. Matolak, S. Kaya, S. Laha, and W. Rayess, "Energy-efficient adaptive wireless nocs architecture," in *2013 Seventh IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*, April 2013, pp. 1–8.

[26] S. Deb, A. Ganguly, P. P. Pande, B. Belzer, and D. Heo, "Wireless noc as interconnection backbone for multicore chips: Promises and challenges," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 2, pp. 228–239, June 2012.

[27] J. J. Lin, H. T. Wu, Y. Su, L. Gao, A. Sugavanam, J. E. Brewer, and K. K. O, "Communication using antennas fabricated in silicon integrated circuits," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 8, pp. 1678–1687, Aug 2007.

[28] M. Zuffada, "The industrialization of the silicon photonics: Technology road map and applications," in *2012 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, Sept 2012, pp. 7–13.

[29] B. Koch, A. Alduino, L. Liao, R. Jones, M. Morse, B. Kim, W. Z. Lo, J. Basak, H. F. Liu, H. Rong, M. Sysak, C. Krause, R. Saba, D. Lazar, L. Horwitz, R. Bar,

S. Litski, A. Liu, K. Sullivan, O. Dosunmu, N. Na, T. Yin, F. Haubensack, I. w. Hsieh, J. Heck, R. Beatty, J. Bovington, and M. Paniccia, "A 12.5 gb/s cwdm si photonics link using integrated hybrid silicon lasers," in *CLEO: 2011 - Laser Science to Photonic Applications*, May 2011, pp. 1–2.

[30] N. Dupuis, "Technologies for fast, scalable silicon photonic switches," in *2015 International Conference on Photonics in Switching (PS)*, Sept 2015, pp. 100–102.

[31] E. Fusella *et al.*, "Lighting up on-chip communications with photonics: Design tradeoffs for optical noc architectures," *IEEE Circuits and Systems Magazine*, vol. 16, no. 3, pp. 4–14, 2016.

[32] G. Li, A. V. Krishnamoorthy, I. Shubin, J. Yao, Y. Luo, H. Thacker, X. Zheng, K. Raj, and J. E. Cunningham, "Ring resonator modulators in silicon for interchip photonic links," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 6, pp. 95–113, Nov 2013.

[33] K. Bergman, L. P. Carloni, A. Biberman, J. Chan, and G. Hendry, *Photonic network-on-chip design.* Springer, 2014.

[34] R. Rodes, J. B. Jensen, D. Zibar, C. Neumeyr, E. Rönneberg, J. Rosskopf, M. Ortsiefer, and I. T. Monroy, "Vertical-cavity surface-emitting laser based digital coherent detection for multigigabit long reach passive optical links," *Microwave and Optical Technology Letters*, vol. 53, no. 11, pp. 2462–2464, 2011.

[35] A. Demir, G. Zhao, S. Freisem, X. Liu, and D. G. Deppe, "Scaling properties of lithographic vcsels," in *SPIE OPTO.* International Society for Optics and Photonics, 2011, pp. 79 520O–79 520O.

[36] E. Macii *et al.*, *Ultra low-power electronics and design.* Springer, 2004.

[37] P. Dong, R. Shafiiha, S. Liao, H. Liang, N.-N. Feng, D. Feng, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, "Wavelength-tunable silicon microring modulator," *Opt. Express*, vol. 18, no. 11, pp. 10 941–10 946, May 2010. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-11-10941

[38] J. Chan, G. Hendry, K. Bergman, and L. P. Carloni, "Physical-layer modeling and system-level design of chip-scale photonic interconnection networks," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 30, no. 10, pp. 1507–1520, 2011.

[39] X. Tu, M. Li, J. Jiang, D. Goodwill, P. Dumais, E. Bernier, H. Fu, and D. Geng, "Compact low-loss adiabatic bends in silicon shallow-etched waveguides," in *2016 IEEE 13th International Conference on Group IV Photonics (GFP)*, Aug 2016, pp. 48–49.

[40] Y. Zhang, S. Yang, A. E. J. Lim, G. Q. Lo, C. Galland, T. Baehr-Jones, and M. Hochberg, “A cmos-compatible, low-loss, and low-crosstalk silicon waveguide crossing,” *IEEE Photonics Technology Letters*, vol. 25, no. 5, pp. 422–425, March 2013.

[41] S. Xiao, M. H. Khan, H. Shen, and M. Qi, “Modeling and measurement of losses in silicon-on-insulator resonators and bends,” *Optics Express*, vol. 15, no. 17, pp. 10 553–10 561, 2007.

[42] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson, “12.5 gbit/s silicon micro-ring silicon modulators,” in *2007 Conference on Lasers and Electro-Optics (CLEO)*, May 2007, pp. 1–2.

[43] I. Kiyat, A. Aydinli, and N. Dagli, “Low-power thermooptical tuning of soi resonator switch,” *IEEE Photonics Technology Letters*, vol. 18, no. 2, pp. 364–366, Jan 2006.

[44] X. Zheng, F. Liu, J. Lexau, D. Patil, G. Li, Y. Luo, H. D. Thacker, I. Shubin, J. Yao, K. Raj, R. Ho, J. E. Cunningham, and A. V. Krishnamoorthy, “Ultralow power 80 gb/s arrayed cmos silicon photonic transceivers for wdm optical links,” *Journal of Lightwave Technology*, vol. 30, no. 4, pp. 641–650, Feb 2012.

[45] H. Gu, J. Xu, and Z. Wang, “Odor: a microresonator-based high-performance low-cost router for optical networks-on-chip,” in *Proceedings of the 6th IEEE/ACM/IFIP international conference on Hardware/Software codesign and system synthesis*. ACM, 2008, pp. 203–208.

[46] F. Xia, M. Rooks, L. Sekaric, and Y. Vlasov, “Ultra-compact silicon wdm optical filters with flat - top response for on-chip optical interconnects,” in *2007 Conference on Lasers and Electro-Optics (CLEO)*, May 2007, pp. 1–2.

[47] G. Kurian, J. E. Miller, J. Psota, J. Eastep, J. Liu, J. Michel, L. C. Kimerling, and A. Agarwal, “Atac: A 1000-core cache-coherent processor with on-chip optical network,” in *2010 19th International Conference on Parallel Architectures and Compilation Techniques (PACT)*, Sept 2010.

[48] J. Chan, G. Hendry, A. Biberman, K. Bergman, and L. P. Carloni, “Phoenixsim: A simulator for physical-layer analysis of chip-scale photonic interconnection networks,” in *2010 Design, Automation Test in Europe Conference Exhibition (DATE 2010)*, March 2010, pp. 691–696.

[49] X. Wang, “Rpnoc: a ring-based packet-switched optical network-on-chip,” *IEEE Photonics Technology Letters*, 2015.

[50] S. L. Beux, J. Trajkovic, I. O'Connor, G. Nicolescu, G. Bois, and P. Paulin, "Optical ring network-on-chip (ornoc): Architecture and design methodology," in *2011 Design, Automation Test in Europe*, March 2011, pp. 1–6.

[51] K. Preston, N. Sherwood-Droz, J. S. Levy, and M. Lipson, "Performance guidelines for wdm interconnects based on silicon microring resonators," in *CLEO: 2011 - Laser Science to Photonic Applications*, May 2011, pp. 1–2.

[52] S. Bartolini, L. Lusnig, and E. Martinelli, "Olympic: A hierarchical all-optical photonic network for low-power chip multiprocessors," in *2013 Euromicro Conference on Digital System Design*, Sept 2013, pp. 56–59.

[53] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, and A. Choudhary, "Firefly: Illuminating future network-on-chip with nanophotonics," in *ACM SIGARCH Computer Architecture News*, vol. 37, no. 3. ACM, 2009, pp. 429–440.

[54] Y. Pan, J. Kim, and G. Memik, "Flexishare: Channel sharing for an energy-efficient nanophotonic crossbar," in *High Performance Computer Architecture (HPCA), 2010 IEEE 16th International Symposium on*. IEEE, 2010, pp. 1–12.

[55] Y. Xu, J. Yang, and R. Melhem, "Channel borrowing: an energy-efficient nanophotonic crossbar architecture with light-weight arbitration," in *Proceedings of the 26th ACM international conference on Supercomputing*. ACM, 2012, pp. 133–142.

[56] A. Shacham, K. Bergman, and L. P. Carloni, "On the design of a photonic network-on-chip," in *First International Symposium on Networks-on-Chip (NOCS'07)*, May 2007, pp. 53–64.

[57] S. Le Beux, H. Li, I. O'Connor, K. Cheshmi, X. Liu, J. Trajkovic, and G. Nicolescu, "Chameleon: Channel efficient optical network-on-chip," in *Proceedings of the conference on Design, Automation & Test in Europe*. European Design and Automation Association, 2014, p. 304.

[58] J. Luo, C. Killian, S. L. Beux, D. Chillet, H. Li, I. OConnor, and O. Sentieys, "Channel allocation protocol for reconfigurable optical network-on-chip," in *2015 Workshop on Exploiting Silicon Photonics for Energy-Efficient High Performance Computing*, Jan 2015, pp. 33–39.

[59] J. Luo, A. Elantably, V. D. Pham, C. Killian, D. Chillet, S. L. Beux, O. Sentieys, and I. O'Connor, "Performance and energy aware wavelength allocation on ring-based wdm 3d optical noc," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2017*, March 2017, pp. 1372–1377.

[60] L. Zhang, X. Tan, M. Yang, and Y. Jiang, “A centralized optical network-on-chip architecture with space-division multiple access,” in *2014 Optical Interconnects Conference*, 2014.

[61] Z. Chen, H. Gu, Y. Yang, and D. Fan, “A hierarchical optical network-on-chip using central-controlled subnet and wavelength assignment,” *Journal of Lightwave Technology*, vol. 32, no. 5, pp. 930–938, 2014.

[62] H. Gu, J. Xu, and W. Zhang, “A low-power fat tree-based optical network-on-chip for multiprocessor system-on-chip,” in *2009 Design, Automation Test in Europe Conference Exhibition*, April 2009, pp. 3–8.

[63] E. Fusella and A. Cilardo, “H2onoc: A hybrid optical-electronic noc based on hybrid topology,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 1, pp. 330–343, Jan 2017.

[64] P. Sanchis, J. V. Galan, A. Brimont, A. Griol, J. Marti, M. A. Piqueras, and J. M. Perdigues, “Low-crosstalk in silicon-on-insulator waveguide crossings with optimized-angle,” in *2007 4th IEEE International Conference on Group IV Photonics*, Sept 2007, pp. 1–3.

[65] H. Jayatilleka, K. Murray, M. Caverley, N. A. F. Jaeger, L. Chrostowski, and S. Shekhar, “Crosstalk in soi microring resonator-based filters,” *Journal of Lightwave Technology*, vol. 34, no. 12, pp. 2886–2896, June 2016.

[66] Y. Xie, M. Nikdast, J. Xu, W. Zhang, Q. Li, X. Wu, Y. Ye, X. Wang, and W. Liu, “Crosstalk noise and bit error rate analysis for optical network-on-chip,” in *Design Automation Conference*, June 2010, pp. 657–660.

[67] M. Nikdast, J. Xu, X. Wu, W. Zhang, Y. Ye, X. Wang, Z. Wang, and Z. Wang, “Systematic analysis of crosstalk noise in folded-torus-based optical networks-on-chip,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 33, no. 3, pp. 437–450, March 2014.

[68] M. Nikdast, J. Xu, L. H. K. Duong, X. Wu, Z. Wang, X. Wang, and Z. Wang, “Fat-tree-based optical interconnection networks under crosstalk noise constraint,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 1, pp. 156–169, Jan 2015.

[69] E. Fusella and A. Cilardo, “Phonocmap: An application mapping tool for photonic networks-on-chip,” in *2016 Design, Automation Test in Europe Conference Exhibition (DATE)*, March 2016, pp. 289–292.

[70] L. H. K. Duong, M. Nikdast, S. L. Beux, J. Xu, X. Wu, Z. Wang, and P. Yang, “A case study of signal-to-noise ratio in ring-based optical networks-on-chip,” *IEEE Design Test*, vol. 31, no. 5, pp. 55–65, Oct 2014.

[71] M. Bahadori, S. Rumley, D. Nikolova, and K. Bergman, “Comprehensive design space exploration of silicon photonic interconnects,” *Journal of Lightwave Technology*, vol. 34, no. 12, June 2016.

[72] S. V. R. Chittamuru and S. Pasricha, “Improving crosstalk resilience with wavelength spacing in photonic crossbar-based network-on-chip architectures,” in *2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug 2015, pp. 1–4.

[73] L. Zhou and A. K. Kodi, “Probe: Prediction-based optical bandwidth scaling for energy-efficient nocs,” in *2013 Seventh IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*, April 2013, pp. 1–8.

[74] C. Li, R. Bai, A. Shafik, E. Z. Tabasy, G. Tang, C. Ma, C. H. Chen, Z. Peng, M. Fiorentino, P. Chiang, and S. Palermo, “A ring-resonator-based silicon photonics transceiver with bias-based wavelength stabilization and adaptive-power-sensitivity receiver,” in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, Feb 2013, pp. 124–125.

[75] C. Chen, J. L. AbellAin, and A. Joshi, “Managing laser power in silicon-photonic noc through cache and noc reconfiguration,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 6, pp. 972–985, June 2015.

[76] C. Chen, T. Zhang, P. Contu, J. Klamkin, A. K. Coskun, and A. Joshi, “Sharing and placement of on-chip laser sources in silicon-photonic nocs,” in *2014 Eighth IEEE/ACM International Symposium on Networks-on-Chip (NoCS)*, Sept 2014, pp. 88–95.

[77] F. Lan, R. Wu, C. Zhang, Y. Pan, and K. t. T. Cheng, “Dlps: Dynamic laser power scaling for optical network-on-chip,” in *2017 22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, Jan 2017, pp. 726–731.

[78] M. Ortín-Obón, L. Ramini, V. Viñals-Yúfera, and D. Bertozzi, “A tool for synthesizing power-efficient and custom-tailored wavelength-routed optical rings,” in *2017 22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, Jan 2017, pp. 300–305.

[79] K. Deb *et al.*, “A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii,” in *International Conference on Parallel Problem Solving From Nature*. Springer, 2000, pp. 849–858.

[80] E. Fusella and A. Cilardo, "Crosstalk-aware mapping for tile-based optical network-on-chip," in *High Performance Computing and Communications (HPCC), 2015 IEEE 7th International Symposium on Cyberspace Safety and Security (CSS), 2015 IEEE 12th International Conferen on Embedded Software and Systems (ICESS), 2015 IEEE 17th International Conference on.* IEEE, 2015, pp. 1139–1142.

[81] D. Q. Binh, N. Q. Huy, and N. H. Hai, "A binary mixed integer coded genetic algorithm for multi-objective optimization of nuclear research reactor fuel reloading," *Kerntechnik*, vol. 79, no. 6, pp. 511–517, 2014.

[82] B. I. Rylander, "Computational complexity and the genetic algorithm," dissertation, University of Idaho Moscow, ID, USA, 2001. [Online]. Available: https://pdfs.semanticscholar.org/7685/a4e47fc5edab04993adc6e57c9de13e242f7.pdf

[83] A. Ghiasi, "Large data centers interconnect bottlenecks," *Opt. Express*, vol. 23, no. 3, pp. 2085–2090, Feb 2015. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-23-3-2085

[84] M. Kennedy and A. K. Kodi, "Laser pooling: Static and dynamic laser power allocation for on-chip optical interconnects," *Journal of Lightwave Technology*, vol. 35, no. 15, pp. 3159–3167, Aug 2017.

[85] Q. Xu, D. Fattal, and R. G. Beausoleil, "Silicon microring resonators with 1.5-um radius," *Opt. Express*, vol. 16, no. 6, pp. 4309–4315, Mar 2008. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-6-4309

[86] Y. Audzevich, P. Watts, S. Kilmurray, and A. W. Moore, "Efficient photonic coding: a considered revision," in *Proceedings of the 2nd ACM SIGCOMM workshop on Green networking.* ACM, 2011, pp. 13–18.

[87] J. L. Abellán, A. K. Coskun, A. Gu, W. Jin, A. Joshi, A. B. Kahng, J. Klamkin, C. Morales, J. Recchio, V. Srinivas, and T. Zhang, "Adaptive tuning of photonic devices in a photonic noc through dynamic workload allocation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 5, pp. 801–814, May 2017.

[88] Z. Li, M. Mohamed, X. Chen, A. Mickelson, and L. Shang, "Device modeling and system simulation of nanophotonic on-chip networks for reliability, power and performance," in *2011 48th ACM/EDAC/IEEE Design Automation Conference (DAC)*, June 2011, pp. 735–740.

[89] Z. Li, M. Mohamed, X. Chen, E. Dudley, K. Meng, L. Shang, A. R. Mickelson, R. Joseph, M. Vachharajani, B. Schwartz, and Y. Sun, "Reliability modeling and management of nanophotonic on-chip networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 1, pp. 98–111, Jan 2012.

[90] S. R. Sarangi, B. Greskamp, R. Teodorescu, J. Nakano, A. Tiwari, and J. Torrellas, "Varius: A model of process variation and resulting timing errors for microarchitects," *IEEE Transactions on Semiconductor Manufacturing*, vol. 21, no. 1, pp. 3–13, Feb 2008.

Personal Publications

- **Jiating Luo**, Cedric Killian, Sebastien Le Beux, Daniel Chillet, Hui Li, Ian O'Connor, and Olivier Sentieys, "Channel allocation protocol for reconfigurable Optical Network-on-Chip," , Workshop on Exploiting Silicon Photonics for energy-efficient high-performance computing (SiPhotonics) at HiPEAC 2015, Amsterdam, Netherlands, January 2015.

- Martha Johanna Sepulveda, Sebastien Le Beux, **Jiating Luo**, Cedric Killian, Daniel Chillet, Hui Li, Ian O'Connor, and Olivier Sentieys, "Communication Aware Design Method for Optical Network-on-Chip" International Symposium on Embedded Multicore/Many-core Systems-on-Chip, MCSoC-15. Politecnico di Torino, Turin, Italy,Turin, Italy. September 2015.

- **Jiating Luo**, Ashraf El Antably, Van dung Pham, Cedric Killian, Daniel Chillet, Sebastien Le Beux, Olivier Sentieys, and Ian O'Connor. " Performance and Energy Aware Wavelength Allocation on Ring-Based WDM 3D Optical NoC". In IEEE/ACM Design Automation and Test in Europe (DATE), page 6, Lausanne, Switzerland, March 2017.

- **Jiating Luo**, Cedric Killian, Sebastien Le Beux, Daniel Chillet, Olivier Sentieys, and Ian O'Connor. " Offline optimization of wavelength allocation and laser power in nanophotonic interconnects". ACM J. Emerg.Technol. Comput. Syst.1, 1, Article 1 (January 2018),19 pages. DOI: 10.1145/3178453.